

; CURRENT APPLICATION NUMBER: US/09/848,834A
 ; CURRENT FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: 60/202,328
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 ; FEATURE:
 ; NAME/KEY: PEPTIDE
 ; LOCATION: (1)..(20)
 ; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium
 ; OTHER INFORMATION: falciparum circumsporozoite (CSP) protein
 US-09-848-834A-3

Query Match 35.8%; Score 95; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DEKIAKMEKASSVFNVNS 36
 DB 1 DEKIAKMEKASSVFNVNS 20

RESULT 12

US-10-239-313A-54
 ; Sequence 54, Application US/10239313A
 ; Publication No. US20030175285A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KLINGUER - HAMOUR, Christine
 ; APPLICANT: CORVAIA, Nathalie
 ; APPLICANT: BECK, Alain
 ; APPLICANT: GOETSCH, Liliane
 ; TITLE OF INVENTION: MOLECULE OF PHARMACEUTICAL INTEREST COMPRISING AT ITS
 ; TITLE OF INVENTION: N-TERMINAL A GLUTAMIC ACID OR A GLUTAMINE IN THE FORM
 ; TITLE OF INVENTION: OF A PHYSIOLOGICALLY ACCEPTABLE STRONG ACID
 ; FILE REFERENCE: 343 727 - US
 ; CURRENT APPLICATION NUMBER: US/10/239,313A
 ; CURRENT FILING DATE: 2002-09-19
 ; PRIOR APPLICATION NUMBER: FR 00/03711
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR APPLICATION NUMBER: PCT 01/70772
 ; PRIOR FILING DATE: 2001-03-22
 ; NUMBER OF SEQ ID NOS: 697
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 54
 ; LENGTH: 19
 ; TYPE: PRT
 ; ORGANISM: Plasmodium malariae
 US-10-239-313A-54

Query Match 33.6%; Score 89; DB 14; Length 19;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNS 36
 DB 1 EKXIAKMEKASSVFNVNS 19

RESULT 13

US-09-932-165-1482
 ; Sequence 1482, Application US/09932165
 ; Publication No. US20030134784A1
 ; GENERAL INFORMATION:
 ; APPLICANT: RAITANO, ARTHUR
 ; APPLICANT: CHALLITA-EID, PIA M.
 ; APPLICANT: PARIS, MARY
 ; APPLICANT: SAFFRAN, DOUGLAS
 ; APPLICANT: AFAR, DANIEL
 ; APPLICANT: LEVIN, ELANA

; APPLICANT: HUBERT, RENE
 ; APPLICANT: GE, WANGMAO
 ; APPLICANT: JAKOBOVITZ, AYA
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND CORRESPONDING PROTEINS ENTITLED
 ; TITLE OF INVENTION: 83P2H3 AND CatIF2B11 USEFUL IN TREATMENT AND
 ; TITLE OF INVENTION: DETECTION OF CANCER
 ; FILE REFERENCE: 51158-20014.00
 ; CURRENT APPLICATION NUMBER: US/09/932,165
 ; CURRENT FILING DATE: 2001-08-17
 ; PRIOR APPLICATION NUMBER: 60/226,329
 ; PRIOR FILING DATE: 2000-08-17
 ; NUMBER OF SEQ ID NOS: 1508
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1482
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 US-09-932-165-1482

Query Match 33.6%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNS 36
 DB 3 EKXIAKMEKASSVFNVNS 21

RESULT 14

US-09-935-384-710
 ; Sequence 710, Application US/09935384
 ; Publication No. US20030166526A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHALLITA-EID, PIA
 ; APPLICANT: HUBERT, RENE
 ; APPLICANT: RAITANO, ARTHUR
 ; APPLICANT: AFAR, DANIEL
 ; APPLICANT: LEVIN, ELANA
 ; APPLICANT: PARIS, MARY
 ; APPLICANT: GE, WANGMAO
 ; APPLICANT: JAKOBOVITZ, AYA
 ; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158P1H4
 ; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
 ; TITLE OF INVENTION: OTHER CANCERS
 ; FILE REFERENCE: 51158-20033.00
 ; CURRENT APPLICATION NUMBER: US/09/935,384
 ; CURRENT FILING DATE: 2001-08-22
 ; PRIOR APPLICATION NUMBER: 60/227,098
 ; PRIOR FILING DATE: 2000-08-22
 ; NUMBER OF SEQ ID NOS: 783
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 710
 ; LENGTH: 21
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 US-09-935-384-710

Query Match 33.6%; Score 89; DB 10; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNS 36
 DB 3 EKXIAKMEKASSVFNVNS 21

RESULT 15

US-09-942-052-711
 ; Sequence 711, Application US/09942052
 ; Publication No. US20030170626A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Paris, Mary

; APPLICANT: Hubert, Rene S.
; APPLICANT: Afar, Daniel
; APPLICANT: Ge, Wangmao
; APPLICANT: Challita-Eid, Pia M.
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 85P1B3
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20028.00
; CURRENT APPLICATION NUMBER: US/09/942,052
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 60/228,432
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 744
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 711
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-942-052-711

Query Match 33.6%; Score 89; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 EKXIAKMEKASSVFNVNS 36
| | | | | | | | | | | | | | | | | | | | | |
Db 3 EKXIAKMEKASSVFNVNS 21

Search completed: March 10, 2004, 10:25:50
Job time : 40.0856 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 15.8755 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-20
Perfect score: 265
Sequence: 1 XHWSGLRPGSGPSLDEKK.....NVVNSSGSPSLHWSGLRFX 51

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR 78: *
1: Piri: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 105 | 39.6 | 388 | 2 A39756 | circumsporozoite p |
| 2 | 99 | 37.4 | 405 | 2 S05428 | circumsporozoite p |
| 3 | 99 | 37.4 | 412 | 1 QZQAF | circumsporozoite p |
| 4 | 99 | 37.4 | 424 | 2 A54533 | circumsporozoite p |
| 5 | 98 | 37.0 | 442 | 2 A54529 | circumsporozoite p |
| 6 | 66 | 24.9 | 332 | 1 OZZQMB | circumsporozoite p |
| 7 | 66 | 24.9 | 348 | 1 OZZQBK | circumsporozoite p |
| 8 | 63 | 23.8 | 264 | 2 A44969 | circumsporozoite p |
| 9 | 63 | 23.8 | 367 | 1 OZZQMY | circumsporozoite p |
| 10 | 62.5 | 23.6 | 82 | 2 I51365 | gonadotropin-relea |
| 11 | 60 | 22.6 | 98 | 2 I50739 | gonadotropin-relea |
| 12 | 59.5 | 22.5 | 315 | 2 F98295 | hypothetical prote |
| 13 | 59.5 | 22.5 | 1613 | 2 JE0272 | low density lipopr |
| 14 | 59.5 | 22.5 | 1613 | 2 JE0273 | low density lipopr |
| 15 | 59 | 22.3 | 429 | 2 A54504 | circumsporozoite p |
| 16 | 59 | 22.3 | 485 | 2 A60610 | circumsporozoite p |
| 17 | 58.5 | 22.1 | 74 | 2 I51092 | gonadotropin relea |
| 18 | 58.5 | 22.1 | 82 | 2 I51355 | gonadotropin relea |
| 19 | 58.5 | 22.1 | 82 | 2 I51331 | gonadotropin relea |
| 20 | 58.5 | 22.1 | 624 | 2 T38006 | probable lysophosp |
| 21 | 58 | 21.9 | 10 | 1 RHPG | gonadoliberin - sh |
| 22 | 58 | 21.9 | 10 | 1 RHPG | gonadoliberin - sh |
| 23 | 58 | 21.9 | 67 | 2 I78541 | gonadoliberin prec |
| 24 | 58 | 21.9 | 89 | 2 I51423 | gonadoliberin prec |
| 25 | 58 | 21.9 | 90 | 1 RHMS | gonadoliberin prec |
| 26 | 58 | 21.9 | 92 | 1 RHUG | gonadoliberin prec |
| 27 | 58 | 21.9 | 92 | 1 RHUG | gonadoliberin prec |
| 28 | 58 | 21.9 | 345 | 2 T29416 | hypothetical prote |
| 29 | 58 | 21.9 | 464 | 2 C70122 | thiophene and fura |

gonadoliberin prec
hemocytin - silkwo
fdHE protein - Esc
affects formate de
affects formate de
conserved hypotet
interleukin-beta
hypothetical prote
interleukin-beta
interleukin-beta
interleukin-1 beta
hypothetical prote
salmon-type gonado
hypothetical prote
probable lysophosp
36K antigen pra -

ALIGNMENTS

RESULT 1

A39756
circumsporozoite protein - Plasmodium reichenowi

C:Species: Plasmodium reichenowi

C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Aug-1999

C:Accession: A39756

R:Bal, A.A.; Goldman, I.P.

J. Biol. Chem. 266, 6686-6689, 1991

A:Title: Circumsporozoite protein gene from Plasmodium reichenowi, a chimpanzee malaria

A:Reference number: A39756; MUID:91201303; PMID:2016283

A:Accession: A39756

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <LAL>

A:Cross-references: GB:M60972; NID:G160228; PIDN:AAA29561.1; PID:G160229

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:312-366/Domain: thrombospondin type 1 repeat homology <THRI>

Query Match 39.6%; Score 105; DB 2; Length 388;

Best Local Similarity 58.5%; Pred. No. 6.7e-05;

Matches 24; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 7 LRPSSGSPD-----EKTIAMKASVFNVNSSG 39

DB 337 IKPSAGKPKDLDYENDLEKKICKMEKCSVFNVNSSG 377

RESULT 2

S05428

circumsporozoite protein - malaria parasite (Plasmodium falciparum) (isolate NF54)

C:Species: Plasmodium falciparum

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jun-2000

C:Accession: S05428; A45527; I60657

R:Campbell, J.R.

Nucleic Acids Res. 17, 5854, 1989

A:Title: DNA sequence of the gene encoding a Plasmodium falciparum malaria candidate va

A:Reference number: S05428; MUID:89345189; PMID:2668895

A:Accession: S05428

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-405 <CAM>

A:Cross-references: EMBL:X15363

R:Caspers, P.; Gentz, R.; Matile, H.; Pink, J.R.; Sinigaglia, F.

Mol. Biochem. Parasitol. 35, 185-190, 1989

A:Title: The circumsporozoite protein gene from NF54, a Plasmodium falciparum isolate u

A:Reference number: A45527; MUID:89364998; PMID:2671723

A:Accession: A45527

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <CAS>

A:Cross-references: GB:M22982; GB:J04650; NID:G160168; PIDN:AAA29527.1; PID:G160169

R:Lockyer, M.J.; Marsh, K.; Newbold, C.I.

```

Mol. Biochem. Parasitol. 37, 275-280, 1989
A:Title: Wild isolates of Plasmodium falciparum show extensive polymorphism in T cell ep
A:Reference number: A60657; MUID:90114334; PMID:2481827
A:Accession: I50657
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 319-336,354-373 <LOC>
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:329-383/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 37.4%; Score 99; DB 2; Length 405;
Best Local Similarity 56.1%; Pred. No. 0.00037; Indels 8; Gaps 1;
Matches 23; Conservative 3; Mismatches 7;

QY 7 LRPGSGPSLD-----EKKIAMEKASSVFVNVSIG 39
DB 354 IKPGSANKPKDLDYNDIEKKICKMEKCSVFVNVSIG 394
:||||:|||||

RESULT 3
OZZQAF
C:Species: Plasmodium falciparum
C:Date: 15-Nov-1994 #sequence_revision 15-Nov-1984 #text_change 09-Jun-2000
C:Accession: A03388
R:Dame, J.B.; Williams, J.L.; McCutchan, T.F.; Weber, J.L.; Wirtz, R.A.; Hochmeyer, W.T.
Science 225, 593-599, 1984
A:Title: Structure of the gene encoding the immunodominant surface antigen on the sporoz
A:Reference number: A03388; MUID:84250215; PMID:6204383
A:Accession: A03388
A:Molecule type: DNA
A:Residues: 1-412 <DAM>
A:Cross-references: GB:K02194; NID:G160160; PIDN:AAA29524.1; PID:G160161
A:Experimental source: clone 708
C:Comment: Residues 1-16 are the probable signal sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:336-390/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 37.4%; Score 99; DB 1; Length 412;
Best Local Similarity 56.1%; Pred. No. 0.00038; Indels 8; Gaps 1;
Matches 23; Conservative 3; Mismatches 7;

QY 7 LRPGSGPSLD-----EKKIAMEKASSVFVNVSIG 39
DB 361 IKPGSANKPKDLDYNDIEKKICKMEKCSVFVNVSIG 401
:||||:|||||

RESULT 4
A54533
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54533
R:del Portillo, H.A.; Nussenzweig, R.S.; Enea, V.
Mol. Biochem. Parasitol. 24, 289-294, 1987
A:Title: Circumsporozoite gene of a Plasmodium falciparum strain from Thailand.
A:Reference number: A54533; MUID:87315205; PMID:3306373
A:Accession: A54533
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <DEI>
A:Cross-references: GB:M19752; NID:G160216; PIDN:AAA29555.1; PID:G160217
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:348-402/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 37.4%; Score 99; DB 2; Length 424;
Best Local Similarity 56.1%; Pred. No. 0.00039; Indels 8; Gaps 1;
Matches 23; Conservative 3; Mismatches 7;

QY 7 LRPGSGPSLD-----EKKIAMEKASSVFVNVSIG 39
DB 373 IKPGSANKPKDLDYNDIEKKICKMEKCSVFVNVSIG 413
:||||:|||||

```

```

RESULT 5
A54529
C:Species: Plasmodium falciparum
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C:Accession: A54529
R:Lockyer, M.J.; Schwarz, R.T.
Mol. Biochem. Parasitol. 22, 101-108, 1987
A:Title: Strain variation in the circumsporozoite protein gene of Plasmodium falciparum
A:Reference number: A54529; MUID:87115616; PMID:3543671
A:Accession: A54529
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-442 <LOC>
A:Cross-references: GB:M15505; NID:G160214; PIDN:AAA29554.1; PID:G160215
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:366-420/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 37.0%; Score 98; DB 2; Length 442;
Best Local Similarity 56.1%; Pred. No. 0.00054; Indels 8; Gaps 1;
Matches 23; Conservative 3; Mismatches 7;

QY 7 LRPGSGPSLD-----EKKIAMEKASSVFVNVSIG 39
DB 391 IKPGSANKPKDLDYNDIEKKICKMEKCSVFVNVSIG 431
:||||:|||||

RESULT 6
OZZOMB
C:Species: Plasmodium berghei
C:Date: 30-Sep-1987 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999
C:Accession: A44948; A25083; S13448
R:Janar, D.E.
Mol. Biochem. Parasitol. 39, 151-154, 1990
A:Title: Sequence of the circumsporozoite gene of Plasmodium berghei ANKA clone and NK6
A:Reference number: A44948; MUID:90158693; PMID:2406593
A:Accession: A44948
A:Molecule type: DNA
A:Residues: 1-332 <LAN>
A:Cross-references: GB:M28887
R:Zehlinger, D.J.; Arnot, D.E.; Tam, J.P.; Nussenzweig, V.; Enea, V.
Mol. Cell. Biol. 6, 3965-3972, 1986
A:Title: Circumsporozoite protein of Plasmodium berghei: gene cloning and identification
A:Reference number: A25083; MUID:87089740; PMID:2432395
A:Accession: A25083
A:Molecule type: DNA
A:Residues: 1-26, 'I', 28-68, 'PMURR', 75-126, 'P', 128-134, 'PPNANDP', 135-332 <EIC>
A:Cross-references: GB:M14135; NID:G160245; PIDN:AAA29577.1; PID:G160246
R:Weber, J.L.; Egan, J.E.; Lyon, J.A.; Wirtz, R.A.; Charoenvit, Y.; Maloy, W.L.; Hockme
Exp. Parasitol. 63, 295-300, 1987
A:Title: Plasmodium berghei: cloning of the circumsporozoite protein gene.
A:Reference number: S13446; MUID:87218962; PMID:3556207
A:Accession: S13446
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 61-122, 'A', 124-332 <WEB>
A:Cross-references: GB:M25445; NID:G160177; PIDN:AAA29531.1; PID:G160178
C:Comment: There are three distinct regions in the mature circumsporozoite protein, the
obc membrane-anchoring sequence.
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-332/Product: circumsporozoite protein #status predicted <MAT>
F:94-189/Region: 8-residue repeats
F:199-230/Region: 2-residue repeats
F:258-310/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 24.9%; Score 66; DB 1; Length 332;

```


A:Reference number: I50739; MUID:95396797; PMID:7667296
A:Accession: I50739
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-38 <WHI>
A:Cross-references: EMBL:U31865; NID:905398; PIDN:AAC59691.1; PID:905399
C:Superfamily: gonadoliberin

Query Match 22.6%; Score 60; DB 2; Length 98;
Best Local Similarity 30.4%; Pred. No. 3.7;
Matches 17; Conservative 6; Mismatches 23; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGPSLDE-----KKIAKMEKASSVFNVNSSGSPSLHWSYGL 48
DB 24 HWSYGLSPGK-RDLNFSDTLGNVVEFPVPEAPCSVFGCAEESPFAKMYRKG 78

RESULT 12
F98295
hypothetical protein AGR_L_2640gl [imported] - Agrobacterium tumefaciens (strain C58, C58)
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C:Accession: F98295
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qureshi, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, D.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743134
A:Accession: F98295
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK9888.1; PID:g15159835; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_2640gl
A:Map position: linear chromosome

Query Match 22.5%; Score 59.5; DB 2; Length 315;
Best Local Similarity 34.5%; Pred. No. 16;
Matches 19; Conservative 4; Mismatches 13; Indels 19; Gaps 2;

QY 6 GLRPGSGPSLDEKKI-----AKMEKASSVFN--VNSSGSPS 41
DB 29 GCRQSGPLLEKEVKRKLFGALALVLAAPALAEATRGFATANNMMSGFS 83

RESULT 13
JE0272
low density lipoprotein receptor-related protein 6 - human
C:Species: Homo sapiens (man)
C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 08-Sep-2002
C:Accession: JE0272
R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metzke
Biochem. Biophys. Res. Commun. 248, 879-888, 1998
A:Title: Isolation and characterization of LRP6, a novel member of the low density lipoprotein receptor family
A:Reference number: JE0272; MUID:98369644; PMID:9704021
A:Accession: JE0272
A:Molecule type: mRNA
A:Residues: 1-1613 <BRO>
A:Cross-references: GB:AF074264; NID:g3462526; PIDN:AAC33006.1; PID:g3462527
C:Genetics:
A:Gene: LRP6
A:Map position: 12p11-12p13
C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL receptor
F:286-323/Domain: EGF homology <EGF1>
F:592-627/Domain: EGF homology <EGF2>
F:1207-1243/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 22.5%; Score 59.5; DB 2; Length 1613;
Best Local Similarity 26.4%; Pred. No. 1e+02;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 18 EKIAKMEKASSVFNVNSSSG 39

Matches 19; Conservative 6; Mismatches 22; Indels 25; Gaps 2;

QY 4 SYGLRPGSGPSLDEKKIAKMEKASSVFN-----VVN-----SSS 38
DB 1448 SLISMGSGSPPYDRAHVTGASSSSSSSTKGTFFAILNPPSPATERSHYTMEFGYSN 1507

QY 39 GPLSHWSYGLRP 50
DB 1508 SPSTHRSYSYRP 1519

RESULT 14
JE0273
low density lipoprotein receptor-related protein 6 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 08-Sep-2002
C:Accession: JE0273
R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metzke
Biochem. Biophys. Res. Commun. 248, 879-888, 1998
A:Title: Isolation and characterization of LRP6, a novel member of the low density lipoprotein receptor family
A:Reference number: JE0272; MUID:98369644; PMID:9704021
A:Accession: JE0273
A:Molecule type: mRNA
A:Residues: 1-1613 <BRO>
A:Cross-references: GB:AF074265; NID:g3462528; PIDN:AAC33007.1; PID:g3462529
C:Genetics:
A:Gene: LRP6
A:Map position: 6
C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL receptor
F:286-323/Domain: EGF homology <EGF1>
F:592-627/Domain: EGF homology <EGF2>
F:1207-1243/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 22.5%; Score 59.5; DB 2; Length 1613;
Best Local Similarity 26.4%; Pred. No. 1e+02; 22; Indels 25; Gaps 2;

Matches 19; Conservative 6; Mismatches 22; Indels 25; Gaps 2;

QY 4 SYGLRPGSGPSLDEKKIAKMEKASSVFN-----VVN-----SSS 38
DB 1448 SLISMGSGSPPYDRAHVTGASSSSSSSTKGTFFAILNPPSPATERSHYTMEFGYSN 1507

QY 39 GPLSHWSYGLRP 50
DB 1508 SPSTHRSYSYRP 1519

RESULT 15
A54504
circumsporozoite protein - Plasmodium malariae (strain Uganda 1-CDC)
C:Species: Plasmodium malariae
C>Date: 06-Oct-1994 #sequence_revision 04-Nov-1994 #text_change 20-Aug-1999
C:Accession: A54504
R:La, A.A.; de la Cruz, V.F.; Campbell, G.H.; Procell, P.M.; Collins, W.E.; McCutchan, M.L. Biochem. Parasitol. 30, 291-294, 1988
A:Title: Structure of the circumsporozoite gene of Plasmodium malariae.
A:Reference number: A54504; MUID:89040027; PMID:3054537
A:Accession: A54504
A:Molecule type: DNA
A:Residues: 1-429 <LAL>
A:Cross-references: GB:J03992; NID:g160220; PIDN:AAA29557.1; PID:g160221
C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology
C:Keywords: tandem repeat
F:354-407/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 22.3%; Score 59; DB 2; Length 429;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 11; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 18 EKIAKMEKASSVFNVNSSSG 39

Db 397 ETEICSLDKCSSIFNVVNSLG 418

Search completed: March 10, 2004, 09:16:53
Job time : 16.9343 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 17.8093 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSGPSLQYIK.....GITELSSGSLHWSYGLRXP 46

Scoring table: ELOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------------|
| 1 | 114.5 | 46.9 | 47 | 1 | US-08-446-692-35 |
| 2 | 114.5 | 46.9 | 47 | 2 | US-08-488-851A-35 |
| 3 | 113.5 | 46.5 | 27 | 1 | US-08-446-692-13 |
| 4 | 113.5 | 46.5 | 27 | 2 | US-08-488-851A-13 |
| 5 | 108.5 | 44.5 | 699 | 2 | US-08-694-865-16 |
| 6 | 108.5 | 44.5 | 699 | 3 | US-09-124-491-16 |
| 7 | 108.5 | 44.5 | 699 | 4 | US-09-383-912-16 |
| 8 | 106 | 43.4 | 49 | 1 | US-08-387-156-4 |
| 9 | 106 | 43.4 | 49 | 2 | US-08-694-865-4 |
| 10 | 106 | 43.4 | 49 | 2 | US-08-878-748-4 |
| 11 | 106 | 43.4 | 49 | 3 | US-09-124-491-4 |
| 12 | 106 | 43.4 | 49 | 4 | US-09-383-912-4 |
| 13 | 106 | 43.4 | 544 | 1 | US-08-387-156-10 |
| 14 | 106 | 43.4 | 544 | 2 | US-08-694-865-10 |
| 15 | 106 | 43.4 | 544 | 2 | US-08-878-748-10 |
| 16 | 106 | 43.4 | 544 | 3 | US-09-124-491-10 |
| 17 | 106 | 43.4 | 544 | 4 | US-09-383-912-10 |
| 18 | 106 | 43.4 | 977 | 1 | US-08-387-156-8 |
| 19 | 106 | 43.4 | 977 | 2 | US-08-694-865-8 |
| 20 | 106 | 43.4 | 977 | 2 | US-08-878-748-8 |
| 21 | 106 | 43.4 | 977 | 3 | US-09-124-491-8 |
| 22 | 106 | 43.4 | 977 | 4 | US-09-383-912-8 |
| 23 | 91 | 37.3 | 40 | 4 | US-09-026-276-35 |
| 24 | 91 | 37.3 | 40 | 4 | US-09-964-201A-35 |
| 25 | 91 | 37.3 | 41 | 4 | US-09-026-276-34 |
| 26 | 91 | 37.3 | 41 | 4 | US-09-964-201A-34 |
| 27 | 90 | 36.9 | 44 | 1 | US-07-690-983D-45 |

Sequence 6, Appl
Sequence 7, Appl
Sequence 47, Appl
Sequence 40, Appl
Sequence 26, Appl
Sequence 30, Appl
Sequence 31, Appl
Sequence 26, Appl
Sequence 29, Appl
Sequence 30, Appl
Sequence 31, Appl
Sequence 31, Appl
Sequence 16, Appl
Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-08-446-692-35
; Sequence 35, Application US/08446692
; Patent No. 5759551

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/446.692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)415-8745

; TELEFAX: (516)751-6849

; INFORMATION FOR SEQ ID NO: 35:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 47 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-446-692-35

Query Match 46.9%; Score 114.5; DB 1; Length 47;

Best Local Similarity 72.7%; Pred. No. 7.4e-08;

Matches 24; Conservative 6; Indels 3; Gaps 1;

QY 13 GPSLOYIKANSKFIGITELSSGSLHWSYGLRP 45

DB 17 GKKOYIKANSKFIGITELGGE--HWSYGLRP 46

```
RESULT 2
US-08-488-351A-35
; Sequence 35, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-35
;
; Query Match 46.9%; Score 114.5; DB 2; Length 47;
; Best Local Similarity 72.7%; Pred. No. 7.4e-08;
; Matches 24; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
;
QY 13 GPSLOVKNKSGFTESSGSPSLHWSYGLRP 45
Db 17 GKKQVKNKSGFTESSGSPSLHWSYGLRP 45
;
RESULT 3
US-08-446-692-13
; Sequence 13, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
```

```
;
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-13
;
; Query Match 46.5%; Score 113.5; DB 1; Length 27;
; Best Local Similarity 79.3%; Pred. No. 5.2e-08;
; Matches 23; Conservative 0; Mismatches 1; Indels 5; Gaps 1;
;
QY 17 QYKANSKFTESSGSPSLHWSYGLRP 45
Db 3 QYKANSKFTESSGSPSLHWSYGLRP 26
;
RESULT 4
US-08-488-351A-13
; Sequence 13, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
```


APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/383,912
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-3400
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-383-912-16
Query Match 44.5%; Score 108.5; DB 4; Length 699;
Best Local Similarity 50.0%; Pred. No. 1e-05;
Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;
QY 2 HWSYGLRPGSGPSLQYIKANSKFGITELSSGSPSLHWSYGLRP 45
DB 36 HWSYGLRPGSGQDSY-----GLRPGSGQHWSYGLRP 68
RESULT 8
US-08-387-156-4
Sequence 4, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUM P.A.
TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-4
Query Match 43.4%; Score 106; DB 1; Length 49;
Best Local Similarity 50.0%; Pred. No. 9.5e-07;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;
QY 2 HWSYGLRPGSGPSLQYIKANSKFGITELSSGSPSLHWSYGLRP 45
DB 2 HWSYGLRPGSGQDSY-----GLRPGSGQHWSYGLRP 35
RESULT 9
US-08-694-865-4
Sequence 4, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-3400
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-4
Query Match 43.4%; Score 106; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 9.5e-07;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-4

Query Match 43.4%; Score 106; DB 4; Length 49;
Best Local Similarity 50.0%; Pred. No. 9.5e-07;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 2 HWSYGLRPGSGQDSY-----GLR--PGSSQHSYGLRP 35

RESULT 13
US-08-387-156-10
; Sequence 10, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid

; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-10

Query Match 43.4%; Score 106; DB 1; Length 544;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 495 HWSYGLRPGSGQDSY-----GLR--PGSSQHSYGLRP 528

RESULT 14
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-10

Query Match 43.4%; Score 106; DB 2; Length 544;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 495 HWSYGLRPGSGQDSY-----GLR--PGSSQHSYGLRP 528

RESULT 15
US-08-878-748-10
; Sequence 10, Application US/08878748
; Patent No. 5963126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-10

Query Match 43.4%; Score 106; DB 2; Length 544;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Caps 2;

Qy 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGPSLHWSYGLRP 45
Db 495 HWSYGLRPGSGSQDWSY-----GLR--FGSSQHWYGLRP 528

Search completed: March 10, 2004, 09:28:56
Job time : 17.8093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 19.358 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274
Sequence: 1 XHWSYGLRPGSSGSLFNNF.....VSASHLEGPLHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------------|
| 1 | 152 | 55.5 | 32 | 1 | US-08-446-692-14 |
| 2 | 152 | 55.5 | 32 | 2 | US-08-488-351A-14 |
| 3 | 120 | 43.8 | 188 | 4 | US-09-396-937-14 |
| 4 | 114 | 41.6 | 31 | 5 | PCT-US93-11703-64 |
| 5 | 114 | 41.6 | 173 | 4 | US-09-396-937-20 |
| 6 | 114 | 41.6 | 452 | 1 | US-07-618-312A-2 |
| 7 | 114 | 41.6 | 452 | 1 | US-07-618-312A-4 |
| 8 | 114 | 41.6 | 452 | 1 | US-08-110-786A-8 |
| 9 | 114 | 41.6 | 452 | 1 | US-08-280-228-2 |
| 10 | 114 | 41.6 | 452 | 1 | US-08-280-228-4 |
| 11 | 114 | 41.6 | 618 | 1 | US-08-668-381A-5 |
| 12 | 114 | 41.6 | 853 | 4 | US-08-913-880C-17 |
| 13 | 114 | 41.6 | 858 | 4 | US-08-913-880C-16 |
| 14 | 114 | 41.6 | 860 | 4 | US-08-913-880C-15 |
| 15 | 114 | 41.6 | 862 | 4 | US-08-913-880C-14 |
| 16 | 114 | 41.6 | 865 | 4 | US-08-913-880C-13 |
| 17 | 114 | 41.6 | 866 | 4 | US-08-913-880C-12 |
| 18 | 114 | 41.6 | 874 | 4 | US-08-913-880C-11 |
| 19 | 114 | 41.6 | 875 | 4 | US-08-913-880C-10 |
| 20 | 114 | 41.6 | 1315 | 4 | US-08-913-880C-1 |
| 21 | 112 | 40.9 | 21 | 1 | US-07-610-525-1 |
| 22 | 112 | 40.9 | 21 | 2 | US-08-661-052-12 |
| 23 | 112 | 40.9 | 21 | 2 | US-08-460-502-8 |
| 24 | 112 | 40.9 | 21 | 2 | US-08-724-774B-5 |
| 25 | 112 | 40.9 | 21 | 3 | US-09-089-595-5 |
| 26 | 112 | 40.9 | 21 | 3 | US-09-382-855-5 |
| 27 | 112 | 40.9 | 21 | 3 | US-09-183-714B-5 |

Sequence 12, Appl
Sequence 10, Appl
Sequence 12, Appl
Sequence 5, Appl
Sequence 12, Appl
Sequence 3, Appl
Sequence 8, Appl
Sequence 5, Appl
Sequence 35, Appl
Sequence 2, Appl
Sequence 66, Appl
Sequence 5, Appl
Sequence 41, Appl
Sequence 8, Appl
Sequence 16, Appl
Sequence 16, Appl

28 112 40.9 21 3 US-09-188-082-12
29 112 40.9 21 3 US-09-171-969-10
30 112 40.9 21 4 US-09-364-088-12
31 112 40.9 21 4 US-09-642-281-5
32 112 40.9 21 4 US-09-102-716-12
33 112 40.9 21 4 US-08-432-483A-3
34 112 40.9 21 4 US-09-148-711A-8
35 112 40.9 21 4 US-09-589-717-5
36 112 40.9 21 4 US-08-945-289-3
37 112 40.9 21 4 US-09-396-937-35
38 112 40.9 21 4 US-09-405-986A-2
39 112 40.9 21 5 PCT-US93-11703-66
40 107 39.1 22 1 US-08-446-692-5
41 107 39.1 22 2 US-08-488-351A-5
42 107 39.1 22 3 US-09-100-409A-41
43 107 39.1 22 5 PCT-US95-13841-8
44 102.5 37.4 699 2 US-08-694-865-16
45 102.5 37.4 699 3 US-09-124-491-16

ALIGNMENTS

RESULT 1
US-08-446-692-14
; Sequence 14, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-14

Query Match 55.5%; Score 152; DB 1; Length 32;
Best Local Similarity 87.9%; Pred. No. 7.7e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 17 FNNFTVSFMLRVPKVSASHLEGPLHWSYGLRP 49
DB 3 FNNFTVSFMLRVPKVSASHLE----HWSYGLRP 31

RESULT 2
US-08-488-351A-14
; Sequence 14, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-8849
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-14
Query Match 55.5%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. No. 7.7e-14;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 17 FNNFTVSVFWRVFKVSASHLEGSLHWSYGLRP 49
Db 3 FNNFTVSVFWRVFKVSASHLE----HWSYGLRP 31
RESULT 3
US-09-396-937-14
; Sequence 14, Application US/09396937
; Patent No. 6645500
; GENERAL INFORMATION:
; APPLICANT: M&E Biotech A/S
; APPLICANT: HALKIER, Torben
; APPLICANT: HAANING, Jesper
; TITLE OF INVENTION: Method for Down-Regulating Osteoprotegerin Ligand

; TITLE OF INVENTION: Activity
; FILE REFERENCE: 22021 PC 1
; CURRENT APPLICATION NUMBER: US/09/396,937
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion of
; OTHER INFORMATION: murine OPG, residues 158-316 modified by
; OTHER INFORMATION: introduction of tetanus toxoid F30 epitope, and
; OTHER INFORMATION: His tag
US-09-396-937-14
Query Match 43.8%; Score 120; DB 4; Length 188;
Best Local Similarity 73.5%; Pred. No. 1.5e-08;
Matches 25; Conservative 2; Mismatches 3; Indels 4; Gaps 1;
QY 11 SSGPSLFNNFTVSVFWRVFKVSASHLEGSLHWS 44
Db 107 SSHNLMFNNFTVSVFWRVFKVSASHLE----NWS 136
RESULT 4
PCT-US93-11703-64
; Sequence 64, Application PC/TUS9311703
; GENERAL INFORMATION:
; APPLICANT: Chiron Mimotopes Pty. Ltd.
; TITLE OF INVENTION: T-Cell Epitopes
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grant B. Green
; STREET: 4560 Horton St.
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11703
; FILING DATE: 28-DEC-1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/984,852
; FILING DATE: 02-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant B.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0222.101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-11703-64
Query Match 41.6%; Score 114; DB 5; Length 31;
Best Local Similarity 95.5%; Pred. No. 1e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 LFNNFTVSVFWRVFKVSASHLE 37

```
Db      :|||||
6 MFNFTVSWLRVPKVSASHLE 27

RESULT 5
US-09-396-937-20
; Sequence 20, Application US/09396937
; Patent No. 6645500
; GENERAL INFORMATION:
; APPLICANT: M&E Biotech A/S
; APPLICANT: HALKIER, Torben
; APPLICANT: HAANING, Jesper
; TITLE OF INVENTION: Method for Down-Regulating Osteoprotegerin Ligand
; FILE REFERENCE: 2021 PC 1
; CURRENT APPLICATION NUMBER: US/09/396,937
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion between
; OTHER INFORMATION: murine OPGL, residues 158-316 with tetanus toxoid
; OTHER INFORMATION: P30 epitope introduced, and His tag
US-09-396-937-20

Query Match      41.6%; Score 114; DB 4; Length 173;
Best Local Similarity 54.5%; Pred. No. 9.6e-06;
Matches 24; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

Qy      17 FNNFTVSWLRVPKVSASHLE-----GPSLHWS 44
|||||
Db      78 FNNFTVSWLRVPKVSASHLEKTSIKIPSSHNLKMGSTKWS 121

RESULT 6
US-07-618-312A-2
; Sequence 2, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-618-312A-4

Query Match      41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 2.8e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-618-312A-2

Query Match      41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 2.8e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      16 LFNNFTVSWLRVPKVSASHLE 37
|||||
Db      83 MFNFTVSWLRVPKVSASHLE 104

RESULT 7
US-07-618-312A-4
; Sequence 4, Application US/07618312A
; Patent No. 5389540
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 14th Floor
; STREET: 2200 Clarendon Boulevard,
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/618,312A
; FILING DATE: 19910516
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford Mr, Arthur R
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 510-51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 0101 703 8750400
; TELEFAX: 0101 703 5253468
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-618-312A-4

Query Match      41.6%; Score 114; DB 1; Length 452;
```

Best Local Similarity 95.5%; Pred. No. 2.8e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

RESULT 8

US-08-110-786A-8
; Sequence 8, Application US/08110786A
; Patent No. 5443966
; GENERAL INFORMATION:
; APPLICANT: FAIRWEATHER, Neil Fraser
; APPLICANT: MAKOFF, Andrew Joseph
; TITLE OF INVENTION: Expression of tetanus toxin fragment C
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye P.C.
; STREET: 1100 No. 5443966th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,786A
; FILING DATE: 23-AUG-1993 1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/777,337
; FILING DATE: 29-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00943
; FILING DATE: 20-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8914122.0
; FILING DATE: 20 June 1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary J. Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-110-786A-8

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 2.8e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

RESULT 9

US-08-280-228-2
; Sequence 2, Application US/08280228
; Patent No. 5571694
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J

; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5571694th Glebe Road
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,228
; FILING DATE: 25-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/618,312
; FILING DATE: 27-NOV-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8926832.0
; FILING DATE: 28-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mary J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-280-228-2

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 2.8e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

RESULT 10

US-08-280-228-4
; Sequence 4, Application US/08280228
; Patent No. 5571694
; GENERAL INFORMATION:
; APPLICANT: Makoff Dr, Andrew J
; APPLICANT: Romanos Dr, Michael A
; APPLICANT: Clare Dr, Jeffrey J
; APPLICANT: Fairweather Dr, Neil F
; TITLE OF INVENTION: VACCINES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5571694th Glebe Road
; CITY: Arlington,
; STATE: Virginia
; COUNTRY: U.S.A.

```
;
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,228
; FILING DATE: 25-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/618,312
; FILING DATE: 27-NOV-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GS 8926532.0
; FILING DATE: 28-NOV-1989
; APPLICATION DATA:
; APPLICATION NUMBER: GB 9006097.1
; FILING DATE: 17-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mary J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-163
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-280-228-4

Query Match 41.6%; Score 114; DB 1; Length 452;
Best Local Similarity 95.5%; Pred. No. 2.8e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 83 MFNNFTVSFWLRVPKVSASHLE 104

RESULT 11
US-08-668-381A-5
; Sequence 5, Application US/08668381A
; Patent No. 5780024
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert H.
; APPLICANT: Fishman, Paul S.
; APPLICANT: Francis, Jonathan W.
; APPLICANT: Hosler, Betsy A.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
; TITLE OF INVENTION: FRAGMENT C HYBRID PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,381A
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 514
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,473
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,164
; REFERENCE/DOCKET NUMBER: 00786/269001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-668-381A-5

Query Match 41.6%; Score 114; DB 1; Length 618;
Best Local Similarity 95.5%; Pred. No. 4.2e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 249 MFNNFTVSFWLRVPKVSASHLE 270

RESULT 12
US-08-913-880C-17
; Sequence 17, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 17
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 463 to 1315 of SEQ ID NO: 1
; US-08-913-880C-17

Query Match 41.6%; Score 114; DB 4; Length 853;
Best Local Similarity 95.5%; Pred. No. 6.2e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSFWLRVPKVSASHLE 37
Db 484 MFNNFTVSFWLRVPKVSASHLE 505

RESULT 13
US-08-913-880C-16
; Sequence 16, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; TITLE OF INVENTION: VACCINE
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 16
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Clostridium tetani
```

```
; FEATURE:
; OTHER INFORMATION: Amino Acids 458 to 1315 of SEQ ID NO: 1
US-08-913-880C-16

Query Match      41.6%; Score 114; DB 4; Length 858;
Best Local Similarity 95.5%; Pred. No. 6.2e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      16 LFNNFTVSFWLRVPKVSASHLE 37
      :|||||:|||||:|||||:|||||:
Db      489 MFNNFTVSFWLRVPKVSASHLE 510

RESULT 14
US-08-913-880C-15
; Sequence 15, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 15
; LENGTH: 860
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 456 to 1315 of SEQ ID NO: 1
US-08-913-880C-15

Query Match      41.6%; Score 114; DB 4; Length 860;
Best Local Similarity 95.5%; Pred. No. 6.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      16 LFNNFTVSFWLRVPKVSASHLE 37
      :|||||:|||||:|||||:|||||:
Db      491 MFNNFTVSFWLRVPKVSASHLE 512

RESULT 15
US-08-913-880C-14
; Sequence 14, Application US/08913880C
; Patent No. 6372225
; GENERAL INFORMATION:
; APPLICANT: MATSUDA, Morihiro
; TITLE OF INVENTION: TETANUS TOXIN FUNCTIONAL FRAGMENT ANTIGEN AND TETANUS
; FILE REFERENCE: 216-380P
; CURRENT APPLICATION NUMBER: US/08/913,880C
; CURRENT FILING DATE: 1997-11-24
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 14
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Clostridium tetani
; FEATURE:
; OTHER INFORMATION: Amino Acids 454 to 1315 of SEQ ID NO: 1
US-08-913-880C-14

Query Match      41.6%; Score 114; DB 4; Length 862;
Best Local Similarity 95.5%; Pred. No. 6.3e-07;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      16 LFNNFTVSFWLRVPKVSASHLE 37
      :|||||:|||||:|||||:|||||:
Db      493 MFNNFTVSFWLRVPKVSASHLE 514

Search completed: March 10, 2004, 09:28:56
Job time : 19.358 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 14.3191 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSGPSLQYIK.....GIPELSGPSLHWSYGLRXP 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 76 | 31-1 | 1315 | 1 BTCLTN | tentoxilysin (EC 3 |
| 2 | 69 | 28-3 | 82 | 2 I51365 | gonadotropin-relea |
| 3 | 65 | 26-6 | 74 | 2 I51092 | gonadotropin relea |
| 4 | 65 | 26-6 | 82 | 2 I51355 | gonadotropin relea |
| 5 | 65 | 26-6 | 82 | 2 I51331 | gonadotropin relea |
| 6 | 61 | 25-0 | 458 | 2 H87624 | peptidase, M23/M37 |
| 7 | 60 | 24-6 | 447 | 2 B83563 | conserved hypotet |
| 8 | 59-5 | 24-4 | 66 | 2 S31029 | gene 84 protein - |
| 9 | 59-5 | 24-4 | 1172 | 2 T00065 | hypothetical prote |
| 10 | 59 | 24-2 | 123 | 2 G48677 | ig heavy chain V-D |
| 11 | 59 | 24-2 | 135 | 2 PH1494 | ig heavy chain V r |
| 12 | 58-5 | 24-0 | 90 | 1 RHMSG | gonadoliberin prec |
| 13 | 58-5 | 24-0 | 374 | 2 E95361 | probable mucinate |
| 14 | 58 | 23-8 | 10 | 1 RHPGG | gonadoliberin - ph |
| 15 | 58 | 23-8 | 10 | 1 RHSHG | gonadoliberin - ph |
| 16 | 58 | 23-8 | 67 | 2 I78541 | gonadoliberin prec |
| 17 | 58 | 23-8 | 89 | 2 I51423 | gonadoliberin prec |
| 18 | 58 | 23-8 | 92 | 1 RHUG | gonadoliberin prec |
| 19 | 58 | 23-8 | 92 | 1 RHRTG | gonadoliberin prec |
| 20 | 57 | 23-4 | 112 | 2 A27887 | ig kappa chain V r |
| 21 | 57 | 23-4 | 112 | 2 C27887 | ig kappa chain V r |
| 22 | 57 | 23-4 | 115 | 2 S38715 | ig kappa chain V r |
| 23 | 57 | 23-4 | 119 | 2 PH1518 | ig heavy chain V r |
| 24 | 57 | 23-4 | 119 | 2 PH1519 | ig heavy chain V r |
| 25 | 57 | 23-4 | 379 | 2 T44656 | homoserine O-acety |
| 26 | 57 | 23-4 | 530 | 2 F84657 | hypothetical prote |
| 27 | 56 | 23-0 | 102 | 2 PH1491 | ig heavy chain V r |
| 28 | 56 | 23-0 | 120 | 2 A49043 | ig kappa chain V r |
| 29 | 56 | 23-0 | 719 | 2 T52510 | hypothetical prote |

| | | | | | | |
|----|----|------|-----|---|--------|-----------------------|
| 30 | 55 | 22-5 | 108 | 2 | E32530 | ig kappa chain V r |
| 31 | 55 | 22-5 | 112 | 2 | D27887 | ig kappa chain V r |
| 32 | 55 | 22-5 | 123 | 2 | F48677 | ig heavy chain V-D |
| 33 | 55 | 22-5 | 249 | 2 | A41497 | 36k antigen pra - |
| 34 | 55 | 22-5 | 266 | 2 | T30913 | acetyltransferase (EC |
| 35 | 55 | 22-5 | 417 | 2 | C82840 | conserved hypotet |
| 36 | 55 | 22-5 | 480 | 2 | F82796 | hypothetical prote |
| 37 | 54 | 22-1 | 10 | 1 | RHAQ1 | gonadoliberin I - |
| 38 | 54 | 22-1 | 92 | 2 | I50644 | gonadoliberin I pr |
| 39 | 54 | 22-1 | 98 | 2 | S33936 | ig kappa chain V r |
| 40 | 54 | 22-1 | 100 | 2 | S26334 | ig light chain V r |
| 41 | 54 | 22-1 | 103 | 2 | PH1034 | ig light chain V r |
| 42 | 54 | 22-1 | 103 | 2 | PH1030 | ig light chain V r |
| 43 | 54 | 22-1 | 103 | 2 | PH1031 | ig light chain V r |
| 44 | 54 | 22-1 | 107 | 2 | D32530 | ig kappa chain V r |
| 45 | 54 | 22-1 | 111 | 2 | PL0257 | ig kappa chain V r |

ALIGNMENTS

RESULT 1

BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N/Alternate names: tetanus neurotoxin
C/Species: Clostridium tetani
C/Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C/Accession: A25689; A25757; A25194; B25194; A60759; S69348; S09364
R/Eisel, U.; Jarasch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.
EMBO J. 5, 2495-2502, 1986
A/Title: Tetanus toxin: Primary structure, expression in E. coli, and homology with bot
A/Reference number: A25689; MUID:87053814; PMID:3536478
A/Accession: A25689
A/Molecule type: DNA
A/Residues: 1-1315 <EIS>
A/Cross-references: GS:X04436; NID:G40769; PIDN:CAA28033.1; PID:G40770
R/Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A/Title: The complete nucleotide sequence of tetanus toxin.
A/Reference number: A25757; MUID:87040747; PMID:3774547
A/Accession: A25757
A/Molecule type: DNA
A/Residues: 1-1315 <FAI>
A/Cross-references: GS:X06214; NID:G40773; PIDN:CAA29564.1; PID:G40774
A/Experimental source: strain CN3911
R/Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A/Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A/Reference number: A25194; MUID:86085672; PMID:3510187
A/Accession: A25194
A/Molecule type: DNA
A/Residues: 743-1315 <FA2>
A/Cross-references: GS:M2739; NID:G144920; PIDN:AAA23282.1; PID:G144921
A/Accession: B25194
A/Molecule type: protein
A/Residues: 865-894 <FA3>
R/Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A/Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal
A/Reference number: A60759; MUID:90035436; PMID:2478476
A/Accession: A60759
A/Molecule type: protein
A/Residues: 461-475 <MAT>
R/Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A/Title: Delination of several DR-restricted tetanus toxin T cell epitopes.
A/Reference number: J50098; MUID:8903918; PMID:2463305
A/Contents: annotation; epitope region
R/Schiavo, G.; Benfenati, F.; Foulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.
Nature 359, 832-835, 1992
A/Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteoly
A/Reference number: S27125; MUID:93063293; PMID:1331807
A/Contents: annotation

R;de Filippis, V.; Vangelista, L.; Schiavo, G.; Tonello, F.; Montecucco, C.
 Eur. J. Biochem. 229, 61-69, 1995
 A;Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
 A;Reference number: S69348; MUID:95262688; PMID:7744050
 A;Accession: S69348
 A;Molecule type: protein
 A;Residues: 2-31 <DEF>
 C;Comment: The precursor of this protein was an extrachromosomal plasmid.
 C;Comment: The precursor is cleaved by endogenous proteinase activity to form light (fra
 dual chains are not toxic when separated). The amino end of the heavy chain (fragment B)
 C;Comment: Fragment B forms ion channels in a lipid bilayer. Fragment C binds to gangli
 C;Comment: This potent neurotoxin binds to peripheral neuronal synapses, is internalized
 presynaptic neurons. It inhibits neurotransmitter release by proteolytic cleavage of sy
 C;Function:
 A;Description: blocks neuroexocytosis via hydrolysis of a Gin-phe peptide bond in synap
 C;Superfamily: tetanus toxin
 C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F;2-457/Product: tentoxylisin light chain (fragment A) #status predicted <TTL>
 F;461-1315/Product: tentoxylisin heavy chain (fragment B.C) #status experimental <TTH>
 F;461-864/Domain: channel forming (fragment B) #status predicted <TXB>
 F;865-1315/Domain: ganglioside binding (fragment C) #status predicted <TXC>
 F;233,237/Binding site: zinc (His) #status predicted
 F;234/Active site: Glu #status predicted

Query Match 31.1%; Score 76; DB 1; Length 1315;
 Best Local Similarity 93.8%; Pred. No. 0.29;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFTGIGTEL 31
 :|||||
 Db 829 MQYIKANSKFTGIGTEL 844
 :|||||

RESULT 2
 I51365
 gonadotropin-releasing hormone - brown trout
 C;Species: Salmo trutta (brown trout)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C;Accession: I51365
 R;Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
 Mol. Cell. Endocrinol. 84, 167-174, 1992
 A;Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.
 A;Reference number: I51040; MUID:92267241; PMID:1587389
 A;Accession: I51365
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-82 <KLU>
 A;Cross-references: EMBL:X79713; NID:G499343; PIDN:CAA56152.1; PID:G499344
 C;Genetics:
 A;Gene: GnRH
 A;Introns: 46/3; 73/3

Query Match 28.3%; Score 69; DB 2; Length 82;
 Best Local Similarity 38.3%; Pred. No. 0.09;
 Matches 18; Conservative 6; Mismatches 19; Indels 4; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
 :|||||
 Db 25 HWSYGLRPGGK-RSVGELEATIKMDTGGVVALPEETSAAHVSELRLP 70
 :|||||

RESULT 3
 I51092
 gonadotropin releasing hormone - chinook salmon (fragment)
 C;Species: Oncorhynchus tshawytscha (chinook salmon)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Aug-1997
 C;Accession: I51092
 R;Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
 Mol. Cell. Endocrinol. 84, 167-174, 1992
 A;Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.
 A;Reference number: I51040; MUID:92267241; PMID:1587389
 A;Accession: I51092
 A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA
 A;Residues: 1-74 <KLU>
 A;Cross-references: EMBL:X79711; NID:G499322; PID:G499323
 C;Genetics:
 A;Gene: GnRH
 A;Introns: 38/3; 65/3

Query Match 26.6%; Score 65; DB 2; Length 74;
 Best Local Similarity 38.3%; Pred. No. 0.26;
 Matches 18; Conservative 5; Mismatches 20; Indels 4; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
 :|||||
 Db 17 HWSYGLRPGGK-RSVGELEATIKMDTGGVVALPEETSAAHVSELRLP 62
 :|||||

RESULT 4
 I51355
 gonadotropin releasing hormone - Atlantic salmon
 C;Species: Salmo salar (Atlantic salmon)
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C;Accession: I51355; I51355
 R;Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
 Mol. Cell. Endocrinol. 84, 167-174, 1992
 A;Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.
 A;Reference number: I51040; MUID:92267241; PMID:1587389
 A;Accession: I51355
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-82 <KLU>
 A;Cross-references: EMBL:X79709; NID:G499341; PID:G499342
 A;Accession: I51355
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-82 <KLU>
 A;Cross-references: EMBL:X74957; NID:G402786; PIDN:CAA52912.1; PID:G402787
 C;Genetics:
 A;Gene: GnRH
 A;Introns: 46/3; 73/3

Query Match 26.6%; Score 65; DB 2; Length 82;
 Best Local Similarity 38.3%; Pred. No. 0.29;
 Matches 18; Conservative 5; Mismatches 20; Indels 4; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
 :|||||
 Db 25 HWSYGLRPGGK-RSVGELEATIKMDTGGVVALPEETSAAHVSELRLP 70
 :|||||

RESULT 5
 I51331
 gonadotropin releasing hormone - brook trout
 C;Species: Salvelinus fontinalis (brook trout)
 C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 05-Nov-1999
 C;Accession: I51331
 R;Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
 Mol. Cell. Endocrinol. 84, 167-174, 1992
 A;Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.
 A;Reference number: I51040; MUID:92267241; PMID:1587389
 A;Accession: I51331
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-82 <KLU>
 A;Cross-references: EMBL:X79712; NID:G499336; PIDN:CAA56151.1; PID:G499337
 C;Genetics:
 A;Gene: GnRH
 A;Introns: 46/3; 73/3

Query Match 26.6%; Score 65; DB 2; Length 82;
 Best Local Similarity 38.3%; Pred. No. 0.29;
 Matches 18; Conservative 5; Mismatches 20; Indels 4; Gaps 2;

QY 2 HWSYGLRPGSGPSLQYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
 :|||||

Db ||||| || | : :: | : | | | | | | | |

25 HWSYGMFLPGKK-RSVGSELEATIKWMDTGGVVALPEETSAHVSRRLRP 70

RESULT 6
H87624
peptidase, M23/M37 family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 10-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87624
B.; Nierman, W.C.; Feildbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
R.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87624
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <STO>
A:Cross-references: GB:AE005673; MID:g13424676; PIDN:AAK24996.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3034

Query Match 25.0%; Score 61; DB 2; Length 458;
Best Local Similarity 31.0%; Pred. No. 6.7;
Matches 13; Conservative 11; Mismatches 14; Indels 4; Gaps 1;

QY 3 WSVGLRPGSGPSLQYIKANSKFITGLSSGPLHSWGIR 44
 :::||||: :::|||::|:

Db 357 YAKGIKEGTVRQGQVV----AYVGTGMSGGPPLHYEVWLK 394

RESULT 7
B83563
conserved hypothetical protein PA0667 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 01-Mar-2002
C:Accession: B83563
R.; Stover, C.K.; Pham, X.-Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Ba
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Lory, S.; Olsen, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83563
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>
A:Cross-references: GB:AE004501; GB:AE004091; NID:g9946537; PIDN:AAG04056.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0667
C:Superfamily: hypothetical protein HI0409

Query Match 24.6%; Score 60; DB 2; Length 447;
Best Local Similarity 36.1%; Pred. No. 8.8;
Matches 13; Conservative 9; Mismatches 10; Indels 4; Gaps 1;

QY 6 GLRPGSGPSLQYIKANSKFITGLSSGPLSHWSY 41
 |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:

Db 363 GIRAGTSVKQQII---GVGMTGLATGPHLHYEF 394

RESULT 8
S31029
gene 84 protein - Mycobacterium phage L5
C:Species: Mycobacterium phage L5
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: S31029
R.; Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
Mol. Microbiol. 7, 407-417, 1993
A>Title: Superinfection immunity of mycobacteriophage L5: applications for genetic trans

QY 2 HWSYGLRPG 10

probable muconate cycloisomerase (EC 5.5.1.1) [imported] - *Sinorhizobium meliloti* (strain ED5361)
C-species: *Sinorhizobium meliloti*

Db 2 HWSYGLRPG 10

RESULT 15

RHSHG

gonadoliberin - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
 C:Accession: A93780; A01411
 R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
 Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
 A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing facto
 A:Reference number: A93780; MUID:72094314; PMID:4550508
 A:Accession: A93780
 A:Molecule type: protein
 A:Residues: 1-10 <BUR>
 A>Note: the natural and synthetic hormones have the same biological activity
 C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl_end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyroglutamic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl_end (Gly) #status experimental

Query Match 23.8%; Score 58; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10

Db 2 HWSYGLRPG 10

Search completed: March 10, 2004, 09:16:52
 Job time : 14.3779 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 8.41245 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-19

Perfect score: 244

Sequence: 1 XHWSYGLRPGSSGSLQYIK.....GITELSSGSPSLHWSYGLRXP 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 76 | 31.1 | 1314 | 1 | TETX_CLOTE |
| 2 | 69 | 28.3 | 82 | 1 | GON3_SALTR |
| 3 | 65 | 26.6 | 74 | 1 | GON3_ONCTS |
| 4 | 65 | 26.6 | 82 | 1 | GON3_SALSA |
| 5 | 64 | 26.2 | 74 | 1 | GON3_ONCWY |
| 6 | 62.5 | 25.6 | 63 | 1 | GON1_MESAU |
| 7 | 60 | 24.6 | 61 | 1 | GON1_SHEEP |
| 8 | 59.5 | 24.4 | 66 | 1 | VGB4_EPMU5 |
| 9 | 59.5 | 24.4 | 90 | 1 | GON1_RANCA |
| 10 | 59.5 | 24.4 | 1259 | 1 | AUT2_HUMAN |
| 11 | 58.5 | 24.0 | 90 | 1 | GON1_MOUSE |
| 12 | 58 | 23.8 | 67 | 1 | GON1_MACMU |
| 13 | 58 | 23.8 | 89 | 1 | GON1_XENLA |
| 14 | 58 | 23.8 | 91 | 1 | GON1_PIG |
| 15 | 58 | 23.8 | 92 | 1 | GON1_HUMAN |
| 16 | 58 | 23.8 | 92 | 1 | GON1_RAT |
| 17 | 58 | 23.8 | 92 | 1 | GON1_TUPGB |
| 18 | 57 | 23.4 | 90 | 1 | GON3_DICLA |
| 19 | 57 | 23.4 | 99 | 1 | GON1_DICLA |
| 20 | 57 | 23.4 | 379 | 1 | MTFX_LEPME |
| 21 | 55 | 22.5 | 249 | 1 | PRA_MYCLE |
| 22 | 55 | 22.5 | 266 | 1 | XINC_CALSA |
| 23 | 55 | 22.5 | 480 | 1 | Y523_XYLFA |
| 24 | 54 | 22.1 | 10 | 1 | GON1_ALLMI |
| 25 | 54 | 22.1 | 92 | 1 | GON1_CHICK |
| 26 | 53.5 | 21.9 | 148 | 1 | HIL_MYTCA |
| 27 | 53.5 | 21.9 | 202 | 1 | HIL_MYTTR |
| 28 | 53.5 | 21.9 | 596 | 1 | SDP_EIMBO |
| 29 | 53.5 | 21.9 | 733 | 1 | VINE_MOUSE |
| 30 | 53.5 | 21.9 | 940 | 1 | MAZ4_SCHCO |
| 31 | 53 | 21.7 | 299 | 1 | YXCX_CYPAA |
| 32 | 53 | 21.7 | 485 | 1 | RT16_MYXXA |
| 33 | 53 | 21.7 | 539 | 1 | RN37_MOUSE |

34 53 21.7 575 1 ACEA_LYCES
35 52.5 21.5 1191 1 LMG2_MOUSE
36 52 21.3 82 1 GON3_ONCWA
37 52 21.3 89 1 GON3_PORNO
38 52 21.3 90 1 GON3_PAGNA
39 52 21.3 90 1 GON3_SPAU
40 52 21.3 94 1 GON1_HAPBU
41 52 21.3 95 1 GON1_MORSA
42 52 21.3 95 1 GON1_PAGMA
43 52 21.3 95 1 GON1_SPAU
44 51.5 21.1 366 1 METX_LEPIN
45 51.5 21.1 493 1 C6AD_DROME

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tetoxylisin) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TETX OR CTP60.
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=87053814; PubMed=3536478;
RA Biesel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
EMBO J. 5:2495-2502(1986).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
Nucleic Acids Res. 14:7809-7812(1986).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / F88; PLASMID=PB88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baumer S., Fricke W.F., Wieser A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RL [4]
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
J. Bacteriol. 165:21-27(1986).
RL [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;
RA Kriegstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";

P49297 lycopersico
Q61092 mus musculus
P30973 oncorhynchus
P51922 porichthys
P51921 pagrus major
P51923 pagrus major
P51918 haplochromi
O73812 morone saxa
P70074 pagrus major
P51919 pagrus major
Q85410 leptospira
Q94409 drosophila

| | | |
|----|---|------|
| DR | InterPro: IPR000395; Peptidase M27. | |
| DR | Pfam: PF01742; Peptidase M27; I. | |
| DR | PRINTS; PS00760; BONTOKILYSIN. | |
| DR | PROSITE; PS00196; Bontoxilysin; 1. | |
| DR | PROSITE; PS00144; ZINC_PROTEASES; 1. | |
| DR | Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid; | |
| KW | 3D-structure; Complete proteome. | |
| FT | INIT MET | 0 |
| FT | CHAIN | 1 |
| FT | CHAIN | 457 |
| FT | METAL | 1314 |
| FT | ACT_SITE | 232 |
| FT | METAL | 233 |
| FT | TRANSMEM | 236 |
| FT | TRANSMEM | 226 |
| FT | DISULFID | 438 |
| FT | DISULFID | 1076 |
| FT | HELIX | 876 |
| FT | TURN | 883 |
| FT | STRAND | 884 |
| FT | TURN | 891 |
| FT | STRAND | 892 |
| FT | STRAND | 894 |
| FT | TURN | 904 |
| FT | TURN | 909 |
| FT | STRAND | 912 |
| FT | STRAND | 915 |
| FT | TURN | 920 |
| FT | TURN | 928 |
| FT | STRAND | 925 |
| FT | STRAND | 932 |
| FT | HELIX | 938 |
| FT | TURN | 941 |
| FT | STRAND | 949 |
| FT | HELIX | 962 |
| FT | TURN | 969 |
| FT | STRAND | 972 |
| FT | STRAND | 980 |
| FT | HELIX | 983 |
| FT | STRAND | 987 |
| FT | TURN | 996 |
| FT | STRAND | 998 |
| FT | TURN | 1006 |
| FT | STRAND | 1010 |
| FT | STRAND | 1020 |
| FT | TURN | 1021 |
| FT | STRAND | 1031 |
| FT | TURN | 1039 |
| FT | STRAND | 1042 |
| FT | TURN | 1048 |
| FT | STRAND | 1050 |
| FT | TURN | 1058 |
| FT | STRAND | 1068 |
| FT | TURN | 1079 |
| FT | STRAND | 1082 |
| FT | HELIX | 1097 |
| FT | TURN | 1106 |
| FT | STRAND | 1112 |
| FT | TURN | 1114 |
| FT | STRAND | 1116 |
| FT | STRAND | 1120 |
| FT | TURN | 1122 |
| FT | TURN | 1123 |
| FT | STRAND | 1127 |
| FT | HELIX | 1132 |
| FT | TURN | 1135 |
| FT | STRAND | 1137 |
| FT | TURN | 1144 |
| FT | STRAND | 1148 |
| FT | TURN | 1155 |
| FT | STRAND | 1159 |
| FT | STRAND | 1163 |
| FT | TURN | 1173 |
| FT | STRAND | 1184 |
| FT | TURN | 1188 |
| FT | STRAND | 1190 |

```

Query Match      31.1%; Score 76; DB 1; Length 1314;
Best Local Similarity 93.8%; Pred. No. 0.074;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LOYIKANSKFIGITEL 31
DB 828 MQYIKANSKFIGITEL 843

RESULT 2
GON3_SALTR STANDARD; PRT; 82 AA.
AC P45653;
DT 01-NOV-1995 (Rel. 32, Created)
DE 10-NOV-1995 (Rel. 32, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Progonadoliberein III precursor (Contains: Gonadoliberein III
DE (Luteinizing hormone releasing hormone III) (Gonadotropin-releasing
DE hormone III) (GnRH-III) (LH-RH_III) (Luliberin III); GnRH-associated
DE Peptide III].
GN GNRH3.
OS Salmo trutta (Brown trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8032;
RN [1]_TaxID=8032;
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Klungland H., Anderson O., Alestroem P.;
RT "The salmon gonadotrophin-releasing hormone encoding gene in
RT salmonids.";
RL Mol. Mar. Biol. Biotechnol. 1:420-425(1992).
DR PIR; I51365; I51365.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X79711; CAA56152.1; -.
DR PIR; I51365; I51365.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 82
FT PEPTIDE 24 33
FT PEPTIDE 27 82
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 82 AA; 9191 MW; 8053E9534A765408 CRC64;

Query Match      28.3%; Score 69; DB 1; Length 82;
Best Local Similarity 38.3%; Pred. No. 0.025;
Matches 18; Conservative 6; Mismatches 19; Indels 4; Gaps 2;

QY 2 HWSYGLRPGSGPSLOYIKANSKFI---GITELSSGPSLHWSYGLRP 45
DB 25 HWSYGLPGGK-RSVGELEATIKMDTGGVVALPEETSAHFSERLRP 70

RESULT 3
GON3_ONCTS STANDARD; PRT; 74 AA.
AC P45653;
DT 01-NOV-1995 (Rel. 32, Created)
DE 10-NOV-1995 (Rel. 32, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Progonadoliberein III precursor (Contains: Gonadoliberein III
DE (Luteinizing hormone releasing hormone III) (Gonadotropin-releasing
DE hormone III) (GnRH-III) (LH-RH_III) (Luliberin III); GnRH-associated
DE Peptide III].
GN GNRH3.
OS Salmo trutta (Brown trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX NCBI_TaxID=8032;
RN [1]_TaxID=8032;
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Klungland H., Anderson O., Alestroem P.;
RT "The salmon gonadotrophin-releasing hormone encoding gene in
RT salmonids.";
RL Mol. Mar. Biol. Biotechnol. 1:420-425(1992).
DR PIR; I51365; I51365.
DR InterPro; IPR002012; GnRH.
DR Pfam; PF00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 82
FT PEPTIDE 24 33
FT PEPTIDE 27 82
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 82 AA; 9191 MW; 8053E9534A765408 CRC64;

Query Match      26.6%; Score 65; DB 1; Length 74;
Best Local Similarity 38.3%; Pred. No. 0.074;
Matches 18; Conservative 5; Mismatches 20; Indels 4; Gaps 2;

QY 2 HWSYGLRPGSGPSLOYIKANSKFI---GITELSSGPSLHWSYGLRP 45
DB 17 HWSYGLPGGK-RSVGELEATIKMDTGGVVALPEETSAHFSERLRP 62

RESULT 4
GON3_SALSA STANDARD; PRT; 82 AA.
AC P35629; P51920;
DT 01-JUN-1994 (Rel. 29, Created)
DE 01-JUN-1994 (Rel. 29, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberein III precursor (Gonadotropin-releasing hormone III)
DE (GnRH-III) (LH-RH_III) (Luliberin III).
GN GNRH3.
OS Salmo salar (Atlantic salmon),
OS Salvelinus fontinalis (Brook trout) (Brook char), and
OS Oncorhynchus nerka (Sockeye salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

```


OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
 OX NCBI_TaxID=8030, 8038, 8023;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC SPECIES=S.salar, and S.fontinalis; TISSUE=Hypothalamus;
 RX MEDLINE=92267241; PubMed=1587389;
 RA Klungland H., Lorens J.B., Andersen O., Kisen G.O., Alestroem P.;
 RT "The Atlantic salmon prepro-gonadotropin releasing hormone gene and
 RT mRNA".
 RL Mol. Cell. Endocrinol. 84:167-174(1992).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC SPECIES=O.nerka; TISSUE=Liver;
 RX MEDLINE=96227617; PubMed=8674859;
 RA Coe I.R., von Schalburg K.R., Sherwood N.M.;
 RT "Characterization of the Pacific salmon gonadotropin-releasing hormone
 RT gene, copy number and transcription start site".
 RL Mol. Cell. Endocrinol. 115:113-122(1995).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC SPECIES=O.nerka; STRAIN=Nikko; TISSUE=Brain;
 RX MEDLINE=96020547; PubMed=8546809;
 RA Ashihara M., Suzuki M., Kubokawa K., Aida K., Urano A.;
 RT "Two differing precursor genes for the salmon-type gonadotropin-
 RT releasing hormone exist in salmonids".
 RL J. Mol. Endocrinol. 15:1-9(1995).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Brain.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X79709; CAA56348.1; -;
 DR EMBL; X74957; CAA52912.1; -;
 DR EMBL; X79712; CAA56151.1; -;
 DR EMBL; X91408; CAA62751.1; -;
 DR EMBL; D31869; BAA06667.1; -;
 DR PIR; I51331; I51331;
 DR PIR; I51356; I51356;
 DR InterPro; IPR002012; GnRH.
 DR Pfam; PF00446; GnRH; 1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Pyroliidone carboxylic acid.
 FT SIGNAL 1 23 BY SIMILARITY.
 FT CHAIN 24 82 PROGNADOLIBERIN III.
 FT PEPTIDE 24 33 GONADOLIBERIN III.
 FT PEPTIDE 37 82 GnRH-ASSOCIATED PEPTIDE III.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID
 FT MOD_RES 33 33 (BY SIMILARITY).
 FT MOD_RES 81 81 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT VARIANT 81 81 H -> Q.
 SQ SEQUENCE 82 AA; 9143 MW; 8053P4E44A765408 CRC64;
 Query Match 26.6%; Score 65; DB 1; Length 82;
 Best Local Similarity 38.3%; Pred. No. 0.084; 20; Indels 4; Gaps 2;
 Matches 18; Conservative 5; Mismatches 5;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
 DB 25 HWSYGLWPGGK-RSVGELEATIKMDTGVVALPEETSAHVSRRLRP 70
 RESULT 5
 GON3_ONCMY STANDARD; PRT; 74 AA.
 ID -GON3_ONCMY

P55246;
 01-OCT-1996 (Rel. 34, Created)
 01-OCT-1996 (Rel. 34, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberein III precursor (Gonadotropin-releasing hormone III)
 DE (GnRH-III) (LH-RH III) (Luliberin III) (Fragment).
 GN GNH3.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RA Klungland H., Anderson O., Alestroem P.;
 RT "The salmon gonadotropin-releasing hormone encoding gene in
 RT salmonids".
 RL Mol. Mar. Biol. Biotechnol. 1:420-425(1992).
 RN [2]
 RN SEQUENCE OF 1-65 FROM N.A.
 RX MEDLINE=93386322; PubMed=1308825;
 RA Alestroem P., Kisen G., Klungland H., Andersen O.;
 RT "Fish gonadotropin-releasing hormone gene and molecular approaches
 RT for control of sexual maturation: development of a transgenic fish
 RT model".
 RL Mol. Mar. Biol. Biotechnol. 1:376-379(1992).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X79710; CAA56149.1; -;
 DR EMBL; S65569; AAD13966.1; -;
 DR InterPro; IPR002012; GnRH.
 DR Pfam; PF00446; GnRH; 1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal; Pyroliidone carboxylic acid.
 FT NON_TER 1 15 BY SIMILARITY.
 FT SIGNAL 16 74 PROGNADOLIBERIN III.
 FT CHAIN 16 74 GONADOLIBERIN III.
 FT PEPTIDE 16 25 GnRH-ASSOCIATED PEPTIDE III.
 FT PEPTIDE 29 74 PYRROLIDONE CARBOXYLIC ACID
 FT MOD_RES 16 16 (BY SIMILARITY).
 FT MOD_RES 25 25 AMIDATION (G-26 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 74 AA; 8254 MW; BD63C46D8228EP84 CRC64;
 Query Match 26.2%; Score 64; DB 1; Length 74;
 Best Local Similarity 38.3%; Pred. No. 0.1; 20; Indels 4; Gaps 2;
 Matches 18; Conservative 5; Mismatches 5;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFI---GITELSSGSPSLHWSYGLRP 45
 DB 17 HWSYGLWPGGK-RSVGELEATIKMDTGVVLPPEETSAHVSRRLRP 62
 RESULT 6
 GON1_MESAU STANDARD; PRT; 63 AA.
 ID -GON1_MESAU
 AC 009163;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberein I precursor [Contains: Gonadoliberein I (LH-RH I)]

(Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I] (Fragment).
 GN GNRH1 OR GNRH OR LHRH.
 OS Mesocricetus auratus (Golden hamster).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
 RN NCBI_TaxID=10036;
 [1]
 RP SEQUENCE FROM N.A.
 RA Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.; Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U91938; AAB51102.1; -.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberin1.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERN1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Pyrrolidone carboxylic acid.
 FT NON_TER 1 1 PROGONADOLIBERIN I.
 FT CHAIN 1 >63 GONADOLIBERIN I.
 FT PEPTIDE 1 10 GnRH-ASSOCIATED PEPTIDE I (BY
 FT PEPTIDE 14 >63 SIMILARITY).
 FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY (BY SIMILARITY).
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 10 10 SIMILARITY).
 FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
 FT NON_TER 63 63 SIMILARITY).
 SQ SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;
 Query Match 25.6%; Score 62.5; DB 1; Length 63;
 Best Local Similarity 31.4%; Pred. No. 0.13; 8; Indels 21; Gaps 3;
 Matches 16; Conservative 6; Mismatches 8;
 QY 2 HWSYGLRPGSGPSLQYIKANSKTFG-----ITELSGGPSL-----HW 39
 |||||
 Db 2 HWSYGLRPGG-----KRNAERLGDSFQEMDXEVDQLAEPQHLECTVHW 44
 RESULT 7
 GONI_SHEEP
 ID GONI_SHEEP STANDARD; PRT; 61 AA.
 AC Q28588;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
 DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I] (Fragment).
 DE GNRH1 OR LHRH.
 GN Ovis aries (Sheep).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC

Bovidae; Caprinae; Ovis.
 RN NCBI_TaxID=9940;
 [1]
 RP SEQUENCE OF 12-61 FROM N.A.
 RC STRAIN=Western range; TISSUE=Hypothalamus;
 RA Rodriguez R.E., Wise M.E.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE OF 1-10.
 RX MEDLINE=72094314; PubMed=4550508;
 RA Burgess R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J., Fellows R., Blackwell R., Vale W., Guillemain R.;
 RA "Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass spectrometry-decapeptide-Edman degradation).";
 RT Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates the secretion of both luteinizing and follicle-stimulating hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J02517; AAA03433.1; -.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberin1.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERN1.
 DR PROSITE; PS00473; GnRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Pyrrolidone carboxylic acid.
 FT NON_TER 1 1
 FT CHAIN 1 >61 PROGONADOLIBERIN I.
 FT PEPTIDE 1 10 GONADOLIBERIN I.
 FT PEPTIDE 14 >61 GnRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
 FT NON_TER 61 61
 SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
 Query Match 24.6%; Score 60; DB 1; Length 61;
 Best Local Similarity 52.2%; Pred. No. 0.27; 2; Indels 8; Gaps 1;
 Matches 12; Conservative 1; Mismatches 2;
 QY 2 HWSYGLRPGSGPSLQYIKANSK 24
 |||||
 Db 2 HWSYGLRPGG-----KRNAK 16
 RESULT 8
 VG84_BPML5
 ID VG84_BPML5 STANDARD; PRT; 66 AA.
 AC Q05301;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE gene 84 protein (GP84).
 DE 84.
 GN Mycobacteriophage L5.
 OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 CC L5-like viruses.
 CC NCBI_TaxID=31757;
 RN [1]
 RP SEQUENCE FROM N.A.

```
RX MEDLINE=93211282; PubMed=9459766;
RA Hatfull G.F., Sarkis G.J.;
RT "DNA sequence, structure and gene expression of mycobacteriophage L5;
RL a phase system for mycobacterial genetics.";
RM Mol. Microbiol. 7:395-405(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z18946; CAA79460.1; -
DR PIR; S31029; S31029.
SQ SEQUENCE 66 AA; 7424 MW; 9C7104C7A4FA74A5 CRC64;
Query Match 24.4%; Score 59.5; DB 1; Length 66;
Best Local Similarity 45.5%; Pred. No. 0.35;
Matches 15; Conservative 2; Mismatches 7; Indels 9; Gaps 2;
QY 5 YGL-----RPGSSGSLQYKANSKFIGITELS 32
DB 36 YGFEVDVYEPGESG-----YIKRNGKFGVTWEVS 64
RESULT 9
GONI_RANCA
ID GONI_RANCA STANDARD; PRT; 90 AA.
AC Q90Y63;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Progonadolibirin I precursor (Contains: Gonadolibirin I (LHRH I)
DE (Luteinizing hormone releasing hormone I) (Gonadotropin releasing
DE hormone I) (GNRH I) (Luliberin I); GNRH-associated peptide I (GAP1)].
GN GNRH1 OR GNRH.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=Forebrain;
RX MEDLINE=21102951; PubMed=11170016;
RA Wang L., Yoo M.S., Kang H.M., Im W.B., Choi H.S., Bogerd J.,
RA Kwon H.B.;
RT "Cloning and characterization of cDNAs encoding the GNRH1 and GNRH2
RT precursors from bullfrog (Rana catesbeiana).";
RL J. Exp. Zool. 289:190-201(2001).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Forebrain.
CC -!- DEVELOPMENTAL STAGE: Expressed at significantly higher levels
CC during post-breedings. Not expressed in pituitary.
CC -!- SIMILARITY: Belongs to the GNRH family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF188754; AAL05972.1; -
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; NAS.
DR GO; GO:0003755; P:hormone mediated signaling; NAS.
DR GO; GO:0000003; P:reproduction; NAS.
DR InterPro; IPR002012; GNRH.
```

```
DR InterPro; IPR004079; Gonadolibirin1.
DR Pfam; PF00446; GNRH; 1.
DR PRINTS; PRO1541; GONADOLIBRN1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Signal;
KW Pyroglutamate carboxylic acid.
KW SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 90 GONADOLIBIRIN I.
FT PEPTIDE 25 34 GONADOLIBIRIN I.
FT PEPTIDE 38 86 GNRH-ASSOCIATED PEPTIDE I (BY
FT SIMILARITY).
FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 34 34 AMIDATION (G-35 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 90 AA; 10291 MW; 317203B4E3DA2FE7 CRC64;
Query Match 24.4%; Score 59.5; DB 1; Length 90;
Best Local Similarity 47.1%; Pred. No. 0.5;
Matches 16; Conservative 4; Mismatches 9; Indels 5; Gaps 3;
QY 2 HWSYGLRPGSSG--PSLQ--YIKANSKFIGITEL 31
DB 26 HWSYGLRPGKREVSLSQSYAEVPE--VSFTEL 58
RESULT 10
AUT2_HUMAN
ID AUT2_HUMAN STANDARD; PRT; 1259 AA.
AC Q8WAX7; Q9Y4F2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Autism susceptibility gene 2 protein.
DE AUTS2 OR KIAA0442.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND VARIANT
SER-303.
RX MEDLINE=22150863; PubMed=12160723;
RA Sultana R., Yu C.-H., Yu J., Munson J., Chen D., Hua W., Estes A.,
RA Cortes F., de la Barria F., Yu D., Haider S.T., Trask B.J., Green E.D.,
RA Raskind W.H., Distche C.M., Wijsman E., Dawson G., Storm D.R.,
RA Schellenberg G.D., Villacres E.C.;
RT "Identification of a novel gene on chromosome 7q11.2 interrupted by a
RT translocation breakpoint in a pair of autistic twins.";
RL Genomics 80:129-129(2002).
RN [2]
SEQUENCE FROM N.A. (ISOFORM 2)
RX MEDLINE=22158633; PubMed=1216954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RN [3]
PRELIMINARY SEQUENCE OF 88-1259 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RN [4]
DNA Res. 4:307-313(1997).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q8WXX7-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q8WXX7-2; Sequence=VSP_003792;
CC
```

CC Note-No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Strongly expressed in brain, skeletal muscle
 CC and kidney. Also expressed in placenta, lung and leukocytes.
 CC -!- DISEASE: This gene is interrupted by a translocation breakpoint in
 CC a pair of autistic twins.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF26917; AAL37411.1; -
 CC DR EMBL; AB007902; BAA23714.2; ALT_INIT.
 CC DR PIR; T00065; T00065.
 CC DR Genew; HGNC:14262; AUTS2.
 CC DR MIM; 607270; -
 CC KW Chromosomal translocation; Polymorphism; Alternative splicing.
 CC FT DOMAIN 383 410 SER-RICH.
 CC FT DOMAIN 525 548 HIS-RICH.
 CC FT DOMAIN 1122 1181 HIS-RICH.
 CC FT VARSPLIC 611 634 Missing (in isoform Short).
 CC FT VARIANT 303 303 /FTID-VSP_003792.
 CC FT VARIANT 303 303 A -> S.
 CC FT CONFLICT 177 177 /FTID-VAR_013864.
 CC FT CONFLICT 177 177 P -> S (IN REF. 2).
 CC SQ SEQUENCE 1259 AA; 138981 MW; A64D17AFF816B591 CRC64;
 CC -----
 CC Query Match 24.4%; Score 59.5; DB 1; Length 1259;
 CC Best Local Similarity 35.9%; Pred. No. 11;
 CC Matches 14; Conservative 6; Mismatches 14; Indels 5; Gaps 1;
 CC -----
 CC QY 6 GLRPGSSGSLQYIKANSKFGITELSSG-----PSLHW 39
 CC DB 1030 GHPMNISTSLDTRMTFMGISPLPGGRPPYPSFW 1068
 CC -----
 CC RESULT 11
 CC GONI_MOUSE STANDARD; PRT; 90 AA.
 CC ID GONI_MOUSE
 CC AC P13562;
 CC DT 01-JAN-1990 (Rel. 13, Created)
 CC DT 01-JAN-1990 (Rel. 13, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
 CC (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 CC hormone I) (GnRH I) (luliberin I); Prolactin release-inhibiting factor
 CC I].
 CC GN GNRH1 OR GNRH.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RP MEDLINE=87069928; PubMed=3024317;
 CC RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 CC Phillips H.S., Nikolics K., Seeburg P.H.;
 CC RT "A deletion truncating the gonadotropin-releasing hormone gene is
 CC responsible for hypogonadism in the tpg mouse.";
 CC RL Science 234:1366-1371 (1986).
 CC CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC CC -!- SUBCELLULAR LOCATION: Secreted.
 CC CC -!- SIMILARITY: Belongs to the GnRH family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M14872; AAA37717.1; -
 CC DR PIR; A47578; RHMSG.
 CC DR MGD; MG1:95789; GnRH.
 CC DR InterPro; IPR002012; GnRH.
 CC DR InterPro; IPR004079; Gonadoliberin1.
 CC DR Pfam; PF00446; GnRH; 1.
 CC DR PRINTS; PR01541; CONADOLIBRN1.
 CC DR PROSITE; PS00473; GnRH; 1.
 CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Placenta; Signal; Pyroglutamate carboxylic acid.
 CC FT SIGNAL 1 21
 CC FT CHAIN 22 90 PROGNADOLIBERIN I.
 CC FT PEPTIDE 22 31 GNADOLIBERIN I.
 CC FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
 CC FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 CC ACTIVITY.
 CC FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
 CC SQ SEQUENCE 90 AA; 10337 MW; 1C0766FAA826E4D9 CRC64;
 CC -----
 CC Query Match 24.0%; Score 58.5; DB 1; Length 90;
 CC Best Local Similarity 27.9%; Pred. No. 0.67; 17; Indels 5; Gaps 1;
 CC Matches 12; Conservative 9; Mismatches 9;
 CC QY 2 HWSYGLRPGSSGSLQYIKANSKFGITELSSGP-----SLHW 39
 CC DB 23 HWSYGLRPGKRNTEHLVSFQEMGKVDQMAEPQHFECTVHW 65
 CC -----
 CC RESULT 12
 CC GONI_MACMU STANDARD; PRT; 67 AA.
 CC ID GONI_MACMU
 CC AC P55247;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
 CC (luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 CC hormone I) (GnRH I) (luliberin I); GnRH-associated peptide I]
 CC (Fragment).
 CC DE GNRH1 OR GNRH OR LHRH.
 CC OS Macaca mulatta (Rhesus macaque).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC OC Cercopithecinae; Macaca.
 CC OX NCBI_TaxID=9544;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RP TISSUE=Hypothalamus;
 CC RX MEDLINE=95124501; PubMed=7545971;
 CC RA Ma Y.J., Costa M.E., Ojeda S.R.;
 CC RT "Developmental expression of the genes encoding transforming growth
 CC factor alpha and its receptor in the hypothalamus of female rhesus
 CC macaques.";
 CC RL Neuroendocrinology 60:346-359 (1994).
 CC CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC CC -!- SUBCELLULAR LOCATION: Secreted.
 CC CC -!- SIMILARITY: Belongs to the GnRH family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

EMBL; S75916; AAB33096.1; -.
DR PIR; I78541;
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GONADOLIBERN1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Signal; Pyrrolidone carboxylic acid.
FT SIGNAL <1 5 BY SIMILARITY.
FT NON_TER 1 1
FT CHAIN 6 >67 PROGNADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT ACT_SITE 19 >67 GnRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 8 6 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
FT NON_TER 67 67
FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
Query Match 23.8%; Score 58; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 7 HWSYGLRPG 15
RESULT 13
GON1_XENLA STANDARD; PRT; 89 AA.
AC P45656;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin I precursor (Gonadotropin-releasing hormone I) (GnRH-I)
(LH-RH) (Luliberin I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a
mammalian-like expression pattern and conserved domains in
GnRH-associated peptide, but brain onset is delayed until
metamorphosis.";
RL Endocrinology 134:1835-1844 (1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; L28040; AAA49728.1; -.
DR PIR; I51423; I51423.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; Gonadoliberin1.
DR Pfam; PF00446; GnRH; 1.

PRINTS; PR01541; GONADOLIBERN1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
Signal; Pyrrolidone carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 89 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
PEPTIDE.
FT PEPTIDE 37 85 GnRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAE0D4284 CRC64;
Query Match 23.8%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
RESULT 14
GON1_PIG STANDARD; PRT; 91 AA.
AC P49921;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
Gonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I).
GN GnRH1 OR GnRH.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RA Weesner G.D., Matteri R.L., Becker B.A.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RX MEDLINE=72114303; PubMed=4946067;
RA Baba Y., Matsuo H., Schally A.V.;
RT "Structure of the porcine LH- and FSH-releasing hormone. II.
Confirmation of the proposed structure by conventional sequential
analyses.";
RL Biochem. Biophys. Res. Commun. 44:459-463 (1971).
RN [3]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE=72065376; PubMed=4942726;
RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
phase method.";
RL Biochem. Biophys. Res. Commun. 45:822-827 (1971).
RN [4]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE=72117544; PubMed=4946275;
RA Baba Y., Arimura A., Schally A.V.;
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
RL Biochem. Biophys. Res. Commun. 45:483-487 (1971).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
the secretion of both luteinizing and follicle-stimulating
hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; L32864; AAA31066.1; -
 InterPro; IPR002012; GnRH.
 InterPro; IPR004079; Gonadoliberin I.
 Pfam; PF00446; GnRH; 1
 PRINTS; PR01541; GONADOLIBERIN.
 PROSITE; PS00473; GnRH; 1.
 Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 Placenta; Signal; Pyrrolidone carboxylic acid.
 SIGNAL 1 23
 CHAIN 24 91
 PEPTIDE 24 33
 ACT_SITE 26 91
 MOD_RES 24 24
 MOD_RES 33 33
 SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 23.8%; Score 58; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.79; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 2 HWSYGLRPG 10
 25 HWSYGLRPG 33

RESULT 15
 GONI HUMAN STANDARD; PRT; 92 AA.
 AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
 (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
 hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
 peptide I].
 OS GnRH OR GnRH OR LHRH.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8936682; PubMed=2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 hormone gene";
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86094338; PubMed=2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 of gonadotropin-releasing hormone and prolactin release-inhibiting
 factor in human and rat";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT SER-16.
 RX MEDLINE=85012739; PubMed=6030951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 releasing hormone";
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE=83126573; PubMed=6760865;

Tan L., Rousseau P.;
 "The chemical identity of the immunoreactive LHRH-like peptide
 biosynthesized in the human placenta";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 RN [5]
 RP VARIANT SER-16.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 Lander E.S.;
 RT "Characterization of single-nucleotide polymorphisms in coding regions
 of human genes";
 RL Nat. Genet. 22:231-238(1999).
 RN [6]
 RP ERRATUM.
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 Lander E.S.;
 RL Nat. Genet. 23:373-373(1999).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 the secretion of both luteinizing and follicle-stimulating
 hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 Lutrepulse or Lutrelief (Ferring Pharmaceuticals) and Relisorm
 (Serono).
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; X01059; CAA25526.1; -
 EMBL; M12578; AAA35916.1; -
 EMBL; X15215; CAA33285.1; -
 PIR; S05308; RHUG.
 DR Genew; HGNC:4419; GnRH1.
 DR MIM; 152760; -
 DR GO; GO:0005625; C:soluble fraction; TAS.
 DR GO; GO:0005183; F:luteinizing hormone-releasing factor activity; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; Gonadoliberin.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERIN.
 DR PROSITE; PS00473; GnRH; 1.
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 DR Placenta; Pharmaceutical; Signal; Polymorphism;
 DR Pyrrolidone carboxylic acid.
 DR SIGNAL 1 23
 CHAIN 24 92
 PEPTIDE 24 33
 ACT_SITE 26 26
 MOD_RES 24 24
 MOD_RES 33 33
 VARIANT 16 16
 SEQUENCE 92 AA; 10380 MW; 30A72221E076FA79 CRC64;
 Query Match 23.8%; Score 58; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

us-09-848-834a-19.open.rsp

Wed Mar 10 10:34:29 2004

Qv 2 HWSYGLRPG 10
| | | | |
Db 25 HWSYGLRPG 33

Search completed: March 10, 2004, 09:13:59
Job time : 9.41245 secs

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|------|----|--------|------------------------------|
| | | Match | % | | | |
| 1 | 76 | 31.1 | 1310 | 2 | Q93N27 | Q93N27 clostridium |
| 2 | 71 | 29.1 | 1331 | 10 | Q9SF03 | Q9SF03 arabidopsis |
| 3 | 71 | 29.1 | 1379 | 10 | Q84R20 | Q84R20 arabidopsis |
| 4 | 67 | 27.5 | 7223 | 16 | Q7URP6 | Q7URP6 rhodospirillum rubrum |
| 5 | 64 | 26.2 | 82 | 13 | Q90VX3 | Q90VX3 oncorhynchus mykiss |
| 6 | 63 | 25.8 | 82 | 13 | Q7I8P9 | Q7I8P9 oncorhynchus mykiss |
| 7 | 62.5 | 25.6 | 329 | 11 | Q7QT24 | Q7QT24 rattus norvegicus |
| 8 | 61 | 25.0 | 458 | 16 | Q9A409 | Q9A409 caulobacter |
| 9 | 61 | 25.0 | 1019 | 10 | Q7XB05 | Q7XB05 viscum album |
| 10 | 60 | 24.6 | 447 | 16 | Q91504 | Q91504 pseudomonas |
| 11 | 59 | 24.2 | 91 | 13 | Q3PRH0 | Q3PRH0 anguilla japonica |
| 12 | 59 | 24.2 | 368 | 5 | Q8N0N3 | Q8N0N3 penaeus monodon |
| 13 | 59 | 24.2 | 385 | 11 | Q9Z201 | Q9Z201 cavia porcellus |
| 14 | 59 | 24.2 | 473 | 2 | Q7VUQ1 | Q7VUQ1 pseudomonas |
| 15 | 59 | 24.2 | 475 | 16 | Q8QQQ3 | Q8QQQ3 pseudomonas |
| 16 | 58.5 | 24.0 | 164 | 5 | Q86D87 | Q86D87 ciona intesa |

| | | | |
|----|----------|----------|--------------------------------------|
| FT | NON TER | 1 | 1 |
| FT | NON TER | 1310 | 1310 |
| SQ | SEQUENCE | 1310 AA; | 9EADDC914418E450 CRC64; |
| | | | 150316 MW; |
| | | | 31.11%; Score 76; DB 2; Length 1310; |

Query Match 31.1%; Score 76; DB 2; Length 1310;

Best Local Similarity 93.8%; Pred. No. 0.59;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LQYKANSKFIGITEL 31
DB 830 MQYKANSKFIGITEL 845

RESULT 2

Q9SF03 Q9SF03 PRELIMINARY; PRT; 1331 AA.
AC Q9SF03;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F26K24.25 protein (Genomic DNA, chromosome 3, P1 clone: MEC18).
GN F26K24.25
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.
OC NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome III BAC F26K24 genomic sequence."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones."

RL DNA Res. 7:217-221 (2000).
DR EMBL; AC016795; AAF23212.1; -
DR EMBL; AP002040; BAB03106.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004871; CPSF_A_C.
DR Pfam; PF03178; CPSF_A; 1.
SQ SEQUENCE 1331 AA; 146497 MW; 2979F40E8975143E CRC64;

Query Match 29.1%; Score 71; DB 10; Length 1331;
Best Local Similarity 36.1%; Pred. No. 2.8;
Matches 13; Conservative 12; Mismatches 9; Indels 2; Gaps 1;

QY 4 SYGLRPGSGPSLQYKANSKFIGI--TELSGSPSL 37
DB 904 SYKLKPGETGKSMELVRVGNHVLVVGTSLSGSPAI 939

RESULT 3

Q84R20 Q84R20 PRELIMINARY; PRT; 1379 AA.
AC Q84R20;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein At3g11960.
GN At3G11960.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.

OX NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RP Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones."
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BR006164; AAP04148.1; -
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004871; CPSF_A_C.
DR InterPro; IPR007087; Znf_G2H2.
DR Pfam; PF03178; CPSF_A; 1_G2H2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 1379 AA; 151853 MW; 451A0B63797F7D92 CRC64;

Query Match 29.1%; Score 71; DB 10; Length 1379;
Best Local Similarity 36.1%; Pred. No. 2.9;
Matches 13; Conservative 12; Mismatches 9; Indels 2; Gaps 1;

QY 4 SYGLRPGSGPSLQYKANSKFIGI--TELSGSPSL 37
DB 952 SYKLKPGETGKSMELVRVGNHVLVVGTSLSGSPAI 987

RESULT 4

Q7URP6 Q7URP6 PRELIMINARY; PRT; 7223 AA.
AC Q7URP6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB5524
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC NCBI_TaxID=117;
RN SEQUENCE FROM N.A.
RP STRAIN=1;
RC MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
RL EMBL; BX294142; CAD74292.1; -
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 7223 AA; 767506 MW; 0031D53A0B28139A CRC64;

Query Match 27.5%; Score 67; DB 16; Length 7223;
Best Local Similarity 42.3%; Pred. No. 69;
Matches 22; Conservative 6; Mismatches 12; Indels 12; Gaps 4;

QY 3 WSYGLRPGSGSP---SLQYIK---ANSKFIGITE-----LSSGSPSLHWSYGL 43
DB 2531 WYFGLTWDSPANGSKYLNPAAGNNQF-GITEVQHADLSNPNSLNSYDL 2581

RESULT 5

Q90VY3 Q90VY3 PRELIMINARY; PRT; 82 AA.
AC Q90VY3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37; I.
 KW Complete proteome.
 SQ SEQUENCE 458 AA; 49491 MW; C477C81FE97BEA24 CRC64;

Query Match 25.0%; Score 61; DB 16; Length 458;
 Best Local Similarity 31.0%; Pred. No. 16;
 Matches 13; Conservative 11; Mismatches 14; Indels 4; Gaps 1;

QY 3 WSYGLRPGSSPSLOYIKANSKFICITELSSGSPSLHWSYGLR 44
 DB 357 YAKGIKPGTRVRQGVV----AYVGSTGMSSGPHLHYEVWLK 394

RESULT 9

Q7XBC5 PRELIMINARY; PRT; 1019 AA.
 ID Q7XBC5
 AC Q7XBC5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Sucrose phosphate synthase.
 OS Viscum album subsp. album.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Santalales; Santalaceae; Viscum.
 OX NCBI_TaxID=104253;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Maenner P., Hamp R.F.;
 RT "Sucrose phosphate synthase from mistletoe (Viscum album ssp.
 album)".
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY331261; AAP94624.1; -;
 SQ SEQUENCE 1019 AA; 114142 MW; 3F81900921FC3D3 CRC64;
 Query Match 25.0%; Score 61; DB 10; Length 1019;
 Best Local Similarity 26.7%; Pred. No. 42;
 Matches 12; Conservative 9; Mismatches 14; Indels 10; Gaps 1;

QY 8 RFGSSGPSLOYIKANSKFICITELSSGSPSLHWSYGLR 42
 DB 182 RSDTGGQKYYVELARALGTMPIGYRVDLTLTQVSAPIHWSYGLR 226

RESULT 10

Q915Q4 PRELIMINARY; PRT; 447 AA.
 ID Q915Q4
 AC Q915Q4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein PA0667.
 GN PA0667.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 15692 / PA01;
 RC MEDLINE=20437337; PubMed=10984043;
 RX Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufragles W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sajer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 DR EMBL; AE004501; AAG04056.1; -;

DR PIR; B83563; B83563.
 DR GO: GO:0004222; P:metalloendopeptidase activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37; I.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 447 AA; 49156 MW; 5E33C252C2C8CD9D CRC64;

Query Match 24.6%; Score 60; DB 16; Length 447;
 Best Local Similarity 36.1%; Pred. No. 21;
 Matches 13; Conservative 9; Mismatches 10; Indels 4; Gaps 1;

QY 6 GLRPGSSGPSLOYIKANSKFICITELSSGSPSLHWSY 41
 DB 363 GIRAGTSVKQGQII----GYVGTGLATGPHLHYEF 394

RESULT 11

Q9PRH0 PRELIMINARY; PRT; 91 AA.
 ID Q9PRH0
 AC Q9PRH0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Prepro-mGnRH precursor (Gonadoliberin) (Gonadotropin-releasing
 hormone) (LH-RH) (Luliberin).
 OS Anguilla japonica (Japanese eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
 OC Anguilla.
 OX NCBI_TaxID=7937;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okubo K., Suetake H., Aida K.;
 RT "A splicing variant for the prepro-mammalian gonadotropin-releasing
 hormone (prepro-mGnRH) mRNA is present in the brain and various
 peripheral tissues of the Japanese eel."
 RL Zool. Sci. 16:645-651 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Okubo K., Suetake H., Aida K.;
 RT "Expression of two gonadotropin-releasing hormone (GnRH) precursor
 genes in various tissues of the Japanese eel and evolution of GnRH.";
 RL Zool. Sci. 16:471-478 (1999).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
 DR EMBL; AB026989; BAA83608.1; -;
 DR EMBL; AB026991; BAA83597.1; -;
 DR GO: GO:0005576; C:extracellular; IEA.
 DR GO: GO:0005183; F:luteinizing hormone-releasing factor activity; IEA.
 DR GO: GO:0007275; P:development; IEA.
 DR InterPro: IPR002012; GnRH.
 DR InterPro: IPR004079; Gonadoliberin.
 DR Pfam: PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERN.
 DR PROSITE; PS00473; GnRH; 1.
 DR Amidation; Hormone; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 32
 FT CHAIN 33 91
 FT SEQUENCE 91 AA; 9893 MW; BA15C9DC08434A7B CRC64;
 Query Match 24.2%; Score 59; DB 13; Length 91;
 Best Local Similarity 66.7%; Pred. No. 4.4;
 Matches 12; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSIQ 17
 DB 24 HWSYGLRPGSGKRGADSLQ 41

```

RESULT 12
Q8N0N3 PRELIMINARY; PRT; 366 AA.
ID Q8N0N3
AC Q8N0N3
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 25, Last annotation update)
DT 01-OCT-2002 (TReMBLrel. 25, Last annotation update)
DE Beta-1,3-glucan binding protein
OS Peneaus monodon (Peneid shrimp)
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Penaeus.
OX NCBI_TaxID=6687;
RN [1]
RP SEQUENCE FROM N.A.
RA Sriyanyaluckana K., Lee S.Y., Soderhall K.;
RT "The beta-1,3-glucan binding protein from the black tiger shrimp,
RT Peneaus monodon."
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF368168; AAM1213.1;
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR008985; ConA-like lectin.
DR InterPro; IPR000757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
SQ SEQUENCE 366 AA; 41497 MW; 6F540A60E83FDD7D CRC64;

Query Match 24.2%; Score 59; DB 5; Length 366;
Best Local Similarity 31.6%; Pred. No. 23;
Matches 12; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHW 39
Db 171 NWFYGLWPGASBGIDILESGRGNDDFGTGLGNQYGGTTLHW 208

RESULT 13
Q9Z201 PRELIMINARY; PRT; 385 AA.
ID Q9Z201
AC Q9Z201
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Fibrinogen A-alpha chain (fragment).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Murakawa M.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF095466; AAC67565.1;
DR HSSP; P02671; 1F2F.
FT NON_TER 1
FT NON_TER 385
SQ SEQUENCE 385 AA; 40909 MW; FA576CA0E2A2101A CRC64;

Query Match 24.2%; Score 59; DB 11; Length 385;
Best Local Similarity 38.1%; Pred. No. 24;
Matches 16; Conservative 2; Mismatches 12; Indels 12; Gaps 2;

QY 2 HWSYGLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHW 43
Db 128 HWI----PSSGSPGSDQSPVS-----PSSGSGHWSGV 157

RESULT 14
Q8VUQ1 PRELIMINARY; PRT; 473 AA.
ID Q8VUQ1
AC Q8VUQ1
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)

```

```

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Toluene resistance protein Trg3.
GN TRG3.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KT2442;
RA Fukumori F., Kishii M.;
RT "Characterization of the toluene-sensitive mutants of Pseudomonas
RT putida KT2442TOL."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB066296; BAB3598.1;
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37; 1.
SQ SEQUENCE 473 AA; 51435 MW; CCB7BFA918C1039 CRC64;

Query Match 24.2%; Score 59; DB 2; Length 473;
Best Local Similarity 38.9%; Pred. No. 31;
Matches 14; Conservative 7; Mismatches 11; Indels 4; Gaps 1;

QY 6 GLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHW 41
Db 390 GIKTGSTVKQGI---GYITGTGLSTGPHLHYEF 421

RESULT 15
Q88QQ3 PRELIMINARY; PRT; 475 AA.
ID Q88QQ3
AC Q88QQ3
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Peptidase, M23/M37 family.
GN PP0435.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Paterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Unterback T., Rizzo M., Lee K., Kosack D., Noesti D., Wedler H.,
RA Lauber J., Stepanovic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016775; AAN6065.1;
DR TIGR; PF0435;
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37; 1.
KW Complete proteome.
SQ SEQUENCE 475 AA; 51663 MW; 9D816AB099E05A31 CRC64;

Query Match 24.2%; Score 59; DB 16; Length 475;
Best Local Similarity 38.9%; Pred. No. 31;
Matches 14; Conservative 7; Mismatches 11; Indels 4; Gaps 1;

QY 6 GLRPGSSGSPSLQYIKANSKFIGITELSSGSPSLHW 41
Db 390 GIKTGSTVKQGI---GYITGTGLSTGPHLHYEF 421

```

DB 392 GIKTGSTVKQOII-----GYIGTTGLSTGPHLHYEF 423

Search completed: March 10, 2004, 09:25:41
Job time : 47.463 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 69.2685 Seconds
(without alignments)
187.635 Million cell updates/sec

Title: US-09-848-834A-19
Perfect score: 244
Sequence: 1 XHWSYGLRPGSGPSQYIK.....GITESSGSLHWSYGLRXP 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 242 | 99.2 | 46 | 5 AAU11430 | AAU11430 Synthetic |
| 2 | 161 | 66.0 | 31 | 5 AAU11426 | AAU11426 Synthetic |
| 3 | 160.5 | 65.8 | 51 | 5 AAU11431 | AAU11431 Synthetic |
| 4 | 147.5 | 60.5 | 47 | 5 AAU11428 | AAU11428 Synthetic |
| 5 | 145 | 59.4 | 50 | 5 AAU11429 | AAU11429 Synthetic |
| 6 | 136 | 55.7 | 28 | 5 AAU11422 | AAU11422 Synthetic |
| 7 | 114.5 | 46.9 | 47 | 2 AAR62723 | AAR62723 LHRH-cont |
| 8 | 113.5 | 46.5 | 27 | 2 AAR62701 | AAR62701 LHRH-cont |
| 9 | 113.5 | 46.5 | 27 | 7 AAD88947 | AAD88947 LHRH pept |
| 10 | 108.5 | 44.5 | 695 | 2 AAU79573 | AAU79573 LKT-GnRH |
| 11 | 108.5 | 44.5 | 695 | 3 AAY58361 | AAY58361 Leukotoxi |
| 12 | 108.5 | 44.5 | 695 | 3 AAY58133 | AAY58133 Gonadotro |
| 13 | 106 | 43.4 | 49 | 2 AAU03944 | AAU03944 GnRH 4-re |
| 14 | 106 | 43.4 | 49 | 2 AAW79567 | AAW79567 GnRH-2. 1 |
| 15 | 106 | 43.4 | 49 | 2 AAW61542 | AAW61542 Peptide h |
| 16 | 106 | 43.4 | 49 | 3 AAY58363 | AAY58363 Four-copy |
| 17 | 106 | 43.4 | 49 | 3 AAY58135 | AAY58135 GnRH anal |
| 18 | 106 | 43.4 | 544 | 2 AAW03943 | AAW03943 LKT-GnRH |
| 19 | 106 | 43.4 | 544 | 2 AAW79570 | AAW79570 LKT-GnRH |
| 20 | 106 | 43.4 | 977 | 2 AAW03942 | AAW03942 LKT-GnRH |
| 21 | 106 | 43.4 | 977 | 2 AAW79569 | AAW79569 LKT-GnRH |
| 22 | 94 | 38.5 | 42 | 3 AAB20865 | AAB20865 GnRH tand |
| 23 | 92.5 | 37.9 | 23 | 3 AAB20864 | AAB20864 GnRH tand |
| 24 | 91 | 37.3 | 30 | 2 AAR07323 | AAR07323 Luteinisi |
| 25 | 91 | 37.3 | 40 | 2 AAY31183 | AAY31183 Ubiquitin |

| | | | | | |
|----|----|------|-----|------------|--------------------|
| 26 | 91 | 37.3 | 40 | 4 AAB71954 | AAB71954 GnRH sequ |
| 27 | 91 | 37.3 | 41 | 2 AAY31182 | AAY31182 Ubiquitin |
| 28 | 91 | 37.3 | 41 | 4 AAB71953 | AAB71953 GnRH sequ |
| 29 | 90 | 36.9 | 40 | 3 AAY96085 | AAY96085 Cattle go |
| 30 | 90 | 36.9 | 283 | 2 AAR11186 | AAR11186 Plasmid p |
| 31 | 90 | 36.9 | 323 | 2 AAR11187 | AAR11187 Plasmid p |
| 32 | 90 | 36.9 | 398 | 3 AAY96090 | AAY96090 BHV-1 tru |
| 33 | 90 | 36.9 | 399 | 3 AAY96093 | AAY96093 BHV-1 tru |
| 34 | 90 | 36.9 | 411 | 3 AAY96089 | AAY96089 GnRH tetr |
| 35 | 90 | 36.9 | 442 | 3 AAY96091 | AAY96091 GnRH tetr |
| 36 | 89 | 36.5 | 34 | 5 AAU11424 | AAU11424 Synthetic |
| 37 | 89 | 36.5 | 36 | 5 AAU11427 | AAU11427 Synthetic |
| 38 | 89 | 36.5 | 216 | 3 AAY92665 | AAY92665 MUC-1 ana |
| 39 | 89 | 36.5 | 750 | 3 AAY92638 | AAY92638 Mutant hu |
| 40 | 88 | 36.1 | 221 | 5 ABB76114 | ABB76114 PEIA-GnRH |
| 41 | 88 | 36.1 | 537 | 7 ABR82481 | ABR82481 Truncated |
| 42 | 88 | 36.1 | 713 | 7 ABR82480 | ABR82480 Modified |
| 43 | 87 | 35.7 | 20 | 2 AAW47438 | AAW47438 Antigenic |
| 44 | 87 | 35.7 | 20 | 2 AAY31174 | AAY31174 Ubiquitin |
| 45 | 87 | 35.7 | 20 | 2 AAY31177 | AAY31177 Ubiquitin |

ALIGNMENTS

RESULT 1
AAU11430
ID AAU11430 standard; peptide; 46 AA.
XX
AC AAU11430;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 11.
XX
KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; GnRH; synthetic immunogen;
KW promiscuous helper T-cell peptide epitope; LHRH; contraceptive;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Clostridium tetani.
OS Mammalia.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1..10
FT /note= "Gonadotrophin releasing hormone epitope (1..10 aa)"
FT Misc-difference 1
FT /label= OTHER
FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
FT Peptide 11..16
FT /note= "Spacer peptide"
FT Peptide 17..31
FT /note= "Tetanus toxoid (830-844 aa)"
FT Peptide 32..37
FT /note= "Spacer peptide"
FT Peptide 38..46
FT /note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
FT Modified-site 46
FT /note= "Amidated glycine or glycineamide"
XX
WO200185763-A2.
XX
PN 15-NOV-2001.
XX
PD 04-MAY-2001; 2001WO-US014363.
XX
PF 05-MAY-2000; 2000US-0202328P.
XX
PR 05-MAY-2000; 2000US-0202328P.
XX


```

FT Peptide 37.42
FT /note= "Spacer peptide"
FT Peptide 43.51
FT /note= "Gonadotrophin releasing hormone epitope (2-10
FT aa)"
FT Modified-site 51
FT /note= "Amidated glycine or glycineamide"
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
XX
XX 05-MAY-2000; 2000US-0202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX
XX Claim 11; Page 12-13; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone, (LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX useful inducing an immune response against GnRH in an animal subject, and
XX as such is useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX
XX Sequence 51 AA;
XX
XX Query Match 65.8%; Score 160.5; DB 5; Length 51;
XX Best Local Similarity 66.0%; Pred. No. 3.9e-13;
XX Matches 33; Conservative 3; Mismatches 7; Indels 7; Gaps 2;
XX
XX QY 2 HWSYGLRPGSSGPSL-----QYIKANSKFCITELSSGSPSLHWSYGLRP 45
XX |||||
XX DB 2 HWSYGLRPGSSGPSLDEKIKAKWEKASSVF-NVNVSSSGPSLHWSYGLRP 50
XX |||||
XX
XX RESULT 4
XX AAU11428
XX ID AAU11428 standard; peptide; 47 AA.
XX
XX AC AAU11429;
XX
XX DT 12-MAR-2002 (first entry)
XX
XX Synthetic immunogen peptide 9.
XX
XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;
XX luteinising hormone releasing hormone; LHRH; contraceptive;
XX promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
XX breast cancer; uterine cancer; gynaecological cancer; endometriosis;
XX uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
XX Plasmodium falciparum.
XX
XX OS Mammalia.
XX
XX OS Synthetic.
XX
XX OS Chimeric.

```

```

XX PH Key Location/Qualifiers
XX FT Peptide 1.10
XX FT /note= "Gonadotrophin releasing hormone epitope (1.10
XX FT aa)"
XX FT Misc-difference 1
XX FT /label= OTHER
XX FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"
XX FT Peptide 11.16
XX FT /note= "Spacer peptide"
XX FT Peptide 17.34
XX FT /note= "Malaria CSP protein (288-302 aa)"
XX FT Peptide 35.38
XX FT /note= "Spacer peptide"
XX FT Peptide 39.47
XX FT /note= "Gonadotrophin releasing hormone epitope (2-10
XX FT aa)"
XX FT Modified-site 47
XX FT /note= "Amidated glycine or glycineamide"
XX
XX WO200185763-A2.
XX
XX 15-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014363.
XX
XX 05-MAY-2000; 2000US-0202328P.
XX
XX (APHT-) APHTON CORP.
XX
XX Grimes S, Michaeli D, Stevens VC;
XX
XX WPI; 2002-049440/06.
XX
XX Novel synthetic immunogen for inducing immune response against
XX gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX
XX Claim 11; Page 11; 43pp; English.
XX
XX The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone, (LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX useful inducing an immune response against GnRH in an animal subject, and
XX as such is useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX
XX SQ Sequence 47 AA;
XX
XX Query Match 60.5%; Score 147.5; DB 5; Length 47;
XX Best Local Similarity 62.5%; Pred. No. 1.6e-11;
XX Matches 30; Conservative 4; Mismatches 7; Indels 7; Gaps 2;
XX
XX QY 2 HWSYGLRPGSSGSPSLQYIKANSKFCIGI-----TELSSGSPSLHWSYGLRP 45
XX |||||
XX DB 2 HWSYGLRPGSSGSPSLKLL---SEIKGVIVHRLEGEVGPSPSLHWSYGLRP 46
XX |||||
XX
XX RESULT 5
XX AAU11429
XX ID AAU11429 standard; peptide; 50 AA.
XX
XX AC AAU11429;
XX
XX DT 12-MAR-2002 (first entry)
XX
XX

```


DE XX Synthetic immunogen peptide 10.

KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;

KW luteinising hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Clostridium tetani.

OS Mammalia.

OS Synthetic.

OS Chimeric.

XX Key

FT Location/Qualifiers

FT Peptide

FT 1..10 /note= "Gonadotropin releasing hormone epitope (1..10 aa)"

FT Misc-difference 1

FT /label= OTHER

FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

FT Peptide

FT 11..16 /note= "Spacer peptide"

FT Peptide

FT 17..37 /note= "Tetanus toxoid (947-967 aa)"

FT Peptide

FT 38..41 /note= "Spacer peptide"

FT Peptide

FT 42..50 /note= "Gonadotropin releasing hormone epitope (2..10 aa)"

FT Modified-site

FT 50 /note= "Amidated glycine or glycineamide"

XX WO200185763-A2.

PN 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.

XX 05-MAY-2000; 2000US-0202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against

PT gonadotropin releasing hormone, comprises fusion peptide having

PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

PT or its analog.

XX Claim 11; Page 11; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific

CC antibodies against gonadotropin releasing hormone (GnRH) also known as

CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide

CC which comprises a promiscuous helper T-cell peptide epitope and

CC immunomimic peptide epitope or its analogue. The synthetic immunogen is

CC as such useful as a contraceptive and in the treatment of diseases

CC such as cancer (of the breast, uterus and other gynaecological cancer),

CC endometriosis, uterine fibroids, benign prostatic hypertrophy and

CC prostate cancer. The immunogen is effective in eliciting high and

CC specific anti-GnRH antibody titres. The present sequence is a synthetic

CC immunogen of the invention

XX Sequence 50 AA;

SQ

Query Match 59.4%; Score 145; DB 5; Length 50;

Best Local Similarity 60.4%; Pred. No. 3.7e-11;

Matches 29; Conservative 6; Mismatches 9; Indels 4; Gaps 1;

OY 2 HWSYGLRPGSGPSLQVKNKXFGITELSS---GPSLHWSYGLRP 45

Db 2 HWSYGLRPGSGPSLQVKNKXFGITELSS---GPSLHWSYGLRP 49

RESULT 6

AAU11422

ID AAU11422 standard; peptide; 28 AA.

XX AC

XX AAU11422;

XX 12-MAR-2002 (first entry)

XX DE

XX Synthetic immunogen peptide 3.

XX Gonadotropin releasing hormone; GnRH; synthetic immunogen;

KW luteinising hormone releasing hormone; LHRH; contraceptive;

KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX Clostridium tetani.

OS Mammalia.

OS Synthetic.

OS Chimeric.

XX Key

FT Location/Qualifiers

FT Peptide

FT 1..15 /note= "Tetanus toxoid sequence (830-844 aa)"

FT Peptide

FT 16..19 /note= "Spacer peptide"

FT Peptide

FT 20..28 /note= "Gonadotropin releasing hormone epitope"

FT Modified-site

FT 29 /note= "Amidated glycine or glycineamide"

XX WO200185763-A2.

PN 15-NOV-2001.

XX 04-MAY-2001; 2001WO-US014363.

XX 05-MAY-2000; 2000US-0202328P.

XX (APHT-) APHTON CORP.

XX Grimes S, Michaeli D, Stevens VC;

XX WPI; 2002-049440/06.

XX Novel synthetic immunogen for inducing immune response against

PT gonadotropin releasing hormone, comprises fusion peptide having

PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope

PT or its analog.

XX Claim 11; Page 8; 43pp; English.

XX The invention relates to a synthetic immunogen for inducing specific

CC antibodies against gonadotropin releasing hormone (GnRH) also known as

CC luteinising hormone releasing hormone (LHRH) comprising a fusion peptide

CC which comprises a promiscuous helper T-cell peptide epitope and

CC immunomimic peptide epitope or its analogue. The synthetic immunogen is

CC as such useful as a contraceptive and in the treatment of diseases

CC such as cancer (of the breast, uterus and other gynaecological cancer),

CC endometriosis, uterine fibroids, benign prostatic hypertrophy and

CC prostate cancer. The immunogen is effective in eliciting high and

CC specific anti-GnRH antibody titres. The present sequence is a synthetic

CC immunogen of the invention

XX Sequence 28 AA;

SQ

Query Match 55.7%; Score 136; DB 5; Length 28;

Best Local Similarity 93.1%; Pred. No. 2.6e-10;

Matches 27; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 17 QYIKANSKFIGITELSSGPHWSYGLRP 45
 Db 1 QYIKANSKFIGITEL--GPSLWSYGLRP 27

RESULT 7
 AAR62723
 ID AAR62723 standard; peptide; 47 AA.
 XX
 AC AAR62723;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; haptan;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..16 "invasin domain"
 FT /note= "invasin domain"
 FT Domain 19..35
 FT /note= "tetanus toxin helper T cell epitope"
 FT Domain 38..47
 FT /note= "LHRH haptan"
 XX
 WO9425060-A1.
 XX
 PD 10-NOV-1994.
 XX
 XX 28-APR-1994; 94WO-US004832.
 XX
 PR 27-APR-1993; 93US-00057166.
 PR 14-APR-1994; 94US-00229275.
 XX
 PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 PI Ladd AE, Wang CY, Zamb T;
 DR WPI; 1994-357910/44.
 XX
 PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
 PT suppress LHRH activity in males and females.
 XX
 PS Claim 8; Page 88; 213pp; English.
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptan containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasin protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 CC domains and between the immune stimulator and haptan components. When the
 CC haptan is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing immunogenic peptide as above which can be used as a potent
 CC vaccine for treating e.g. prostatic hyperplasia, androgen-dependent
 CC carcinoma, prostatic carcinoma, testicular carcinoma, endometriosis,
 CC benign uterine tumours, recurrent functional ovarian cysts, (severe)
 CC premenstrual syndrome or oestrogen-dependent breast cancer, or for
 CC induction of infertility. (Updated on 25-MAR-2003 to correct FN field.)
 XX

SQ Sequence 47 AA;
 Query Match 46.9%; Score 114.5; DB 2; Length 47;
 Best Local Similarity 72.7%; Pred. No. 2.8e-07;
 Matches 24; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 13 GPSLQYIKANSKFIGITELSSGPHWSYGLRP 45
 Db 17 GGGKQYIKANSKFIGITELGGE---HWSYGLRP 46

RESULT 8
 AAR62701
 ID AAR62701 standard; peptide; 27 AA.
 XX
 AC AAR62701;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-SEP-1995 (first entry)
 XX
 DE LHRH-containing immunogenic peptide.
 XX
 KW Helper T cell epitope; universal immune stimulator; invasin; haptan;
 KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
 KW androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..17
 FT /note= "tetanus toxin helper T cell epitope"
 FT Domain 18..27
 FT /note= "LHRH haptan"
 XX
 WO9425060-A1.
 XX
 PD 10-NOV-1994.
 XX
 XX 28-APR-1994; 94WO-US004832.
 XX
 PR 27-APR-1993; 93US-00057166.
 PR 14-APR-1994; 94US-00229275.
 XX
 PA (LADD/) LADD A E.
 PA (WANG/) WANG C Y.
 PA (ZAMB/) ZAMB T.
 PI Ladd AE, Wang CY, Zamb T;
 DR WPI; 1994-357910/44.
 XX
 PT Immunogenic luteinising hormone releasing hormone peptide(s) - that
 PT suppress LHRH activity in males and females.
 XX
 PS Claim 8, 12; Page 84; 213pp; English.
 CC Synthetic immunogenic peptides are provided in which a universal immune
 CC stimulator is linked to a peptide or protein haptan containing B cell
 CC and/or cytotoxic T lymphocyte epitopes, giving a product which causes
 CC potent immune responses to the coupled peptide or protein. The stimulator
 CC consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
 CC immune response to the coupled peptide in members of a heterogeneous
 CC population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
 CC sequence from the invasin protein of Yersinia. Spacer amino acid
 CC sequences (e.g. Gly-Gly) can be provided between the invasin and Th
 CC domains and between the immune stimulator and haptan components. When the
 CC haptan is LHRH, then optionally the invasin domain can be omitted from
 CC the immune stimulator component. The present sequence represents an LHRH-
 CC containing immunogenic peptide as above which can be used
 CC as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
 CC dependent carcinoma, prostatic carcinoma, testicular carcinoma,
 CC endometriosis, benign uterine tumours, recurrent functional ovarian
 CC cysts, (severe) premenstrual syndrome or oestrogen-dependent breast

KW Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;
 KW immunocontraception; vaccine; feline; canine; equine; cervine.
 XX Mammalia.
 OS Mannheimia haemolytica.
 OS Chimeric.
 XX WO9962545-A2.
 XX 09-DEC-1999.
 XX 28-MAY-1999; 99WO-CA000493.
 XX 04-JUN-1998; 98US-0088024P.
 XX 06-MAY-1999; 99US-00306689.
 XX (BIOS-) BIOSTAR INC.
 XX Robbins SC;
 XX WPI; 2000-086857/07.
 XX N-PSDB; AAZ55700.
 XX Hormone immunogens, analogues or antibodies used to manufacture vaccines
 PT for suppression of reproductive behavior and fertility in vertebrates.
 XX Claim 20; Fig 6A-6F; 88pp; English.
 XX This sequence represents a fusion protein comprising gonadotropin-
 CC releasing hormone (GnRH) immunogens and a Pasteurella haemolytica
 CC leukotoxin (LKT) protein. The fusion protein comprises, in the N to C-
 CC terminal direction, a synthetic peptide sequence (AAV58364), an eight
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC sequence of AAV58363), the LKT protein (which functions as a carrier
 CC protein), and a second eight copy GnRH multimer. The fusion protein may
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies that
 CC cross-react with endogenous GnRH of a vertebrate subject are used to
 CC manufacture a composition or vaccine for immunosterilisation or
 CC immunocontraception of feline, canine, equine or cervine subjects. The
 CC vaccines are used to suppress reproductive behaviour and/or fertility for
 CC at least 10 months. The prepubertal administration results in a
 CC prolonged, long-term suppression of testicular development and/or
 CC function in males, or a prolonged, long-term suppression of ovarian
 CC development and/or function in females. The methods provide a viable and
 CC desirable alternative to surgical forms of sterilisation that are
 CC currently used. (Updated on 12-SEP-2003 to standardise OS field)
 XX Sequence 695 AA;
 SQ Query Match 44.5%; Score 108.5; DB 3; Length 695;
 Best Local Similarity 50.0%; Pred. No. 4.3e-05;
 Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSLHWSYGLRP 45
 DB 36 HWSYGLRPGSGQDSY-----GLRPGSGHWSYGLRP 68
 RESULT 12
 ID AAY58133
 XX AAY58133 standard; protein; 695 AA.
 XX AAY58133;
 XX 12-SEP-2003 (revised)
 DT 07-MAR-2000 (first entry)
 XX Gonadotropin releasing hormone-leukotoxin fusion protein.
 XX Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
 KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
 KW

KW non-androgenic; steroid; reduction; weight gain; muscle distribution;
 KW fat distribution; male pattern; boar taint; flavour; impairment;
 XX reliable; immunocastration; meat production.
 XX Mammalia.
 OS Mannheimia haemolytica.
 OS Chimeric.
 XX WO9956771-A2.
 XX 11-NOV-1999.
 XX 05-MAY-1999; 99WO-CA000360.
 XX 05-MAY-1998; 98US-0084217P.
 XX (BIOS-) BIOSTAR INC.
 XX Manns JG, Acres SD, Harland R;
 XX WPI; 2000-062125/05.
 XX N-PSDB; AA246400.
 XX Production of uncastrated male food animals using vaccines.
 XX Claim 22; Fig 3A-3F; 87pp; English.
 XX This sequence represents a chimeric gonadotropin releasing hormone (GnRH)
 CC -leukotoxin (LKT) fusion protein, which may be used as a vaccine. The LKT
 CC portion of the protein acts to enhance the immunogenicity of the
 CC multimeric GnRH portion (AAV58135). The invention relates to a method of
 CC using two GnRH immunogen vaccines to produce uncastrated male animals for
 CC meat production, one vaccination prior to or during the fattening period
 CC to reduce circulating testosterone levels, and the second vaccination
 CC about 2-8 weeks before slaughter to substantially reduce androgenic
 CC and/or non-androgenic steroids. The invention is used to produce food
 CC animals that exhibit the weight gain and muscle/fat distribution of male
 CC animals without the problems associated with male animals. Such problems
 CC include "boar taint", a urine-like odour found in cooked meat of
 CC uncastrated pigs which is caused by steroids stored in the tissues, and
 CC similar flavour impairments in the meat of other intact male animals. The
 CC invention is more reliable than prior art immunocastration techniques.
 CC (Updated on 12-SEP-2003 to standardise OS field)
 XX Sequence 695 AA;
 SQ Query Match 44.5%; Score 108.5; DB 3; Length 695;
 Best Local Similarity 50.0%; Pred. No. 4.3e-05;
 Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;
 QY 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSSGSLHWSYGLRP 45
 DB 36 HWSYGLRPGSGQDSY-----GLRPGSGHWSYGLRP 68
 RESULT 13
 ID AAW03944
 XX AAW03944 standard; protein; 49 AA.
 XX AAW03944;
 XX 20-NOV-1996 (first entry)
 DT GnRH 4-repeat sequence.
 XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH; fusion protein;
 KW immunogen; vaccine; fertility control; contraceptive; sterilisation;
 KW PCB113; PCB111.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Peptide 1..10


```

XX (BIOS-) BIOSTAR INC.
XX PA
XX PI Harland R, Manns JG, Acres SD;
XX DR WPI; 1998-445952/38.
XX DR N-PSDB; AAV45190.
XX
XX Immunisation against endogenous molecules by administering vaccine to ear
XX PT - useful to elicit efficient and uniform immune response against e.g.
XX PT gonadotrophin releasing hormone to immunocastrate pigs and cattle.
XX PS Example 1; Fig 1B; 61pp; English.
XX
XX This represents the amino acid sequence of the gonadotrophin releasing
XX hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
XX GnRH polypeptide gene fusions. This is used to exemplify the method of
XX invention of immunisation against endogenous molecules by administering a
XX vaccine which comprises an immunogen and a carrier to the ear of the
XX mammal. The method is useful for eliciting an efficient and uniform
XX immune response to block or suppress the activity of an endogenous
XX hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
XX or to elicit an immune response against a targeted endogenous cell type
XX (e.g. a cancerous or otherwise diseased cell). It is especially useful to
XX reduce the levels of GnRH in domestic animals, especially in porcine or
XX bovine species. The use of GnRH immunogens in the vaccine reduces the
XX levels of luteinizing hormone and follicle stimulating hormone and helps
XX in immunocastrating the animal. Administration of vaccine compositions to
XX the ear instead of intramuscular administration into the neck increases
XX the efficiency of vaccination of mammals against endogenous immunogens,
XX and may increase uniformity of vaccine presentation since the ear is
XX relatively uniform from animal to animal
XX SQ Sequence 49 AA;

```

```

Query Match 43.4%; Score 106; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 3.6e-06;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;
OY 2 HWSYGLRPGSSGSPSLQYIKANSKFGITELSSGSPSLHWSYGLRP 45
Db . ||||| ||||| | : ||||| |||||
2 HWSYGLRPGSSGSPSLQYIKANSKFGITELSSGSPSLHWSYGLRP 35

```

Search completed: March 10, 2004, 09:12:14
Job time : 70.2685 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: March 10, 2004, 09:16:59 ; Search time 36.1556 Seconds
(without alignments)
268.645 Million cell updates/sec

Title: US-09-848-834A-19
Perfect score: 244
Sequence: 1 XHWSYGLRPGSSGSLQYIK.....GITELSGPSLHWSYGLRXP 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|----|-------------------|
| 1 | 242 | 99.2 | 46 | 9 | US-09-848-834A-19 |
| 2 | 161 | 66.0 | 31 | 9 | US-09-848-834A-15 |
| 3 | 160.5 | 65.8 | 51 | 9 | US-09-848-834A-20 |
| 4 | 147.5 | 60.5 | 47 | 9 | US-09-848-834A-17 |
| 5 | 145 | 59.4 | 50 | 9 | US-09-848-834A-18 |
| 6 | 136 | 55.7 | 28 | 9 | US-09-848-834A-11 |
| 7 | 113.5 | 46.5 | 27 | 14 | US-10-076-674-7 |
| 8 | 113.5 | 46.5 | 27 | 15 | US-10-355-161A-7 |
| 9 | 108.5 | 44.5 | 695 | 10 | US-09-305-924-13 |
| 10 | 106 | 43.4 | 49 | 9 | US-09-019-010-4 |
| 11 | 106 | 43.4 | 49 | 10 | US-09-305-924-11 |
| 12 | 91 | 37.3 | 40 | 10 | US-09-964-201A-35 |
| 13 | 91 | 37.3 | 41 | 10 | US-09-964-201A-34 |
| 14 | 89 | 36.5 | 34 | 9 | US-09-848-834A-13 |
| 15 | 89 | 36.5 | 36 | 9 | US-09-848-834A-16 |

Sequence 26, Appl
Sequence 29, Appl
Sequence 30, Appl
Sequence 31, Appl
Sequence 14, Appl
Sequence 12, Appl
Sequence 14, Appl
Sequence 8, Appl
Sequence 3, Appl
Sequence 15, Appl
Sequence 4, Appl
Sequence 46, Appl
Sequence 47, Appl
Sequence 2, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 6, Appl
Sequence 9, Appl
Sequence 177, App
Sequence 145, App
Sequence 143, App
Sequence 147, App
Sequence 149, App
Sequence 141, App
Sequence 145, App
Sequence 1, Appl
Sequence 2, Appl
Sequence 4, Appl

20 10 US-09-964-201A-26
20 10 US-09-964-201A-29
20 10 US-09-964-201A-30
20 10 US-09-964-201A-31
37 9 US-09-848-834A-14
33 9 US-09-848-834A-12
158 14 US-10-297-942-14
158 14 US-10-297-942-8
137 14 US-10-297-942-4
287 14 US-10-295-074-15
158 14 US-10-295-074-4
194 14 US-10-295-074-46
194 14 US-10-295-074-47
16 9 US-09-848-834A-2
17 15 US-10-346-563-23
17 15 US-10-321-717-23
17 15 US-09-943-548-8
50 9 US-10-339-522-8
158 14 US-10-297-942-6
285 14 US-10-295-074-9
573 15 US-10-452-024-177
872 14 US-10-241-596-145
879 14 US-10-241-596-143
887 14 US-10-241-596-147
1310 15 US-10-452-024-149
1315 14 US-10-241-596-141
1315 15 US-10-452-024-145
31 14 US-10-237-656-1
15 9 US-09-862-849-2
15 9 US-09-785-215-4

ALIGNMENTS

RESULT 1

US-09-848-834A-19
; Sequence 19, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 19
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetan
; OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino acid
; OTHER INFORMATION: sequence 1-10 of GnRH
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD RES
; LOCATION: (46)..(46)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor

```
; OTHER INFORMATION: (Tentoxylisin)
; NAME/KEY: PEPTIDE
; LOCATION: (32)..(37)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (38)..(46)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19

Query Match          99.2%; Score 242; DB 9; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.4e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
DB 2 HWSYGLRPGSSGSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45

RESULT 2
US-09-848-834A-15
; Sequence 15, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
; OTHER INFORMATION: RH hormone linked by a spacer to amino sequence 830-844 of the Te
; OTHER INFORMATION: tanus toxoid precursor (Tentoxylisin)
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(31)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
US-09-848-834A-15

Query Match          66.0%; Score 161; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.5e-13;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSGSLQYIKANSKFIGITEL 31
DB 2 HWSYGLRPGSSGSLQYIKANSKFIGITEL 31

RESULT 3
US-09-848-834A-20
; Sequence 20, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
```

```
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plasmic
; OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
; LOCATION: (51)..(51)
; OTHER INFORMATION: Amidated glycine or glycylamide
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
; NAME/KEY: PEPTIDE
; LOCATION: (11)..(16)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (17)..(36)
; OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
; OTHER INFORMATION: circumsporozoite (CSP) protein
; NAME/KEY: PEPTIDE
; LOCATION: (37)..(42)
; OTHER INFORMATION: Spacer peptide
; NAME/KEY: PEPTIDE
; LOCATION: (43)..(51)
; OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match          65.8%; Score 160.5; DB 9; Length 51;
Best Local Similarity 66.0%; Pred. No. 5e-13;
Matches 33; Conservative 3; Mismatches 7; Indels 7; Gaps 2;

QY 2 HWSYGLRPGSSGSL-----QYIKANSKFIGITELSSGSPSLHWSYGLRP 45
DB 2 HWSYGLRPGSSGSLDEKKIAXKXKSSVF-NVWNSSSGSPSLHWSYGLRP 50

RESULT 4
US-09-848-834A-17
; Sequence 17, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the (
; OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
; OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
; OTHER INFORMATION: uence 1-10 of the GnRH hormone
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
; NAME/KEY: MOD_RES
```


LOCATION: (47)..(47)
OTHER INFORMATION: Amidated-glycine or glycineamide
NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
NAME/KEY: PEPTIDE
LOCATION: (11)..(18)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (19)..(34)
OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
NAME/KEY: PEPTIDE
LOCATION: (35)..(38)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (39)..(47)
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

Query Match 60.5%; Score 147.5; DB 9; Length 47;
Best Local Similarity 62.5%; Pred. No. 2e-11; 7; Indels 7; Gaps 2;
Matches 30; Conservative 4; Mismatches 7

QY 2 HWSYGLRPGSSGSPSLQYIKANSKFTGI---TELSSGSPSLHWSYGLRP 45
||||| : : : : :
DB 2 HWSYGLRPGSSGSPSLKLL--SEIKGVIVHRLEGVGGPSLHWSYGLRP 46
||||| : : : : :
TYPE: PRT
LENGTH: 50

RESULT 5
US-09-848-834A-18
; Sequence 18, Application US/09848834A
; Patent No. US20020076416A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
; OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 947-967 of the Tet
; OTHER INFORMATION: anus toxoid precursor (Tentoxylisin) protein linked by a spacer b
; OTHER INFORMATION: o amino acid sequence 2-10 of human GnRH
NAME/KEY: MOD RES
LOCATION: (1)..(1)
OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
NAME/KEY: MOD RES
LOCATION: (50)..(50)
OTHER INFORMATION: Amidated glycine or glycineamide
NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
NAME/KEY: PEPTIDE
LOCATION: (11)..(16)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (17)..(37)
OTHER INFORMATION: Amino acid sequence 947-967 of the Tetanus toxoid precursor (Tent
; OTHER INFORMATION: oxylysin
NAME/KEY: PEPTIDE
LOCATION: (38)..(41)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (42)..(50)
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone

US-09-848-834A-18
Query Match 59.4%; Score 145; DB 9; Length 50;
Best Local Similarity 60.4%; Pred. No. 4.4e-11;
Matches 29; Conservative 6; Mismatches 9; Indels 4; Gaps 1;
QY 2 HWSYGLRPGSSGSPSLQYIKANSKFTGITELSS----GPSLHWSYGLRP 45
||||| : : : : :
DB 2 HWSYGLRPGSSGSPSLFNNFTVSFLRVPKVSASHLEGPSLHWSYGLRP 49
||||| : : : : :
RESULT 6
US-09-848-834A-11
; Sequence 11, Application US/09848834A
; Patent No. US20020076418A1
; GENERAL INFORMATION:
; APPLICANT: Aphton Corporation
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047
; CURRENT APPLICATION NUMBER: US/09/848,834A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 830-844 of tl
; OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to a
; OTHER INFORMATION: ino acid sequence 2-10 of the GnRH hormone
NAME/KEY: MOD RES
LOCATION: (1)..(1)
OTHER INFORMATION: Amidated-glutamine
NAME/KEY: MOD RES
LOCATION: (28)..(28)
OTHER INFORMATION: Amidated-glycine or glycineamide
NAME/KEY: PEPTIDE
LOCATION: (1)..(15)
OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus Toxoid Precursor
; OTHER INFORMATION: (Tentoxylisin)
NAME/KEY: PEPTIDE
LOCATION: (16)..(19)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (20)..(28)
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-11

Query Match 55.7%; Score 136; DB 9; Length 28;
Best Local Similarity 93.1%; Pred. No. 3.1e-10;
Matches 27; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
QY 17 QYIKANSKFTGITELSSGSPSLHWSYGLRP 45
||||| : : : : :
DB 1 QYIKANSKFTGITEL--GPSLHWSYGLRP 27
||||| : : : : :
RESULT 7
US-10-076-674-7
; Sequence 7, Application US/10076674
; Publication No. US20030165478A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7

```

; LENGTH: 27
; TYPE: PRT
; ORGANISM: Human
US-10-076-674-7

Query Match 46.5%; Score 113.5; DB 14; Length 27;
Best Local Similarity 79.3%; Pred. No. 2e-07;
Matches 23; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 17 QYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 3 QYIKANSKFIGITELE-----HWSYGLRP 26

RESULT 8
US-10-355-161A-7
; Sequence 7, Application US/10355161A
; Publication No. US2004009897A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Human
US-10-355-161A-7

Query Match 46.5%; Score 113.5; DB 15; Length 27;
Best Local Similarity 79.3%; Pred. No. 2e-07;
Matches 23; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 17 QYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 3 QYIKANSKFIGITELE-----HWSYGLRP 26

RESULT 9
US-09-305-924-13
; Sequence 13, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: GRNH
US-09-305-924-13

Query Match 44.5%; Score 108.5; DB 10; Length 695;
Best Local Similarity 50.0%; Pred. No. 3.3e-05;
Matches 22; Conservative 0; Mismatches 11; Indels 11; Gaps 1;

Qy 2 HWSYGLRPGSGSPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 36 HWSYGLRPGSGSQDWSY-----GLRPGSGSHWSYGLRP 68

RESULT 10
US-09-019-010-4
; Sequence 4, Application US/09019010
; Patent No. US20010014330A1
; GENERAL INFORMATION:
; APPLICANT: HARLAND, RICHARD
; APPLICANT: MANN, JOHN G.
; APPLICANT: ACRES, STEPHEN D.
; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,010
; FILING DATE: 05-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,883
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, Thomas P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-010-4

Query Match 43.4%; Score 106; DB 9; Length 49;
Best Local Similarity 50.0%; Pred. No. 3.5e-06;
Matches 22; Conservative 1; Mismatches 11; Indels 10; Gaps 2;

Qy 2 HWSYGLRPGSGSPSLQYIKANSKFIGITELSSGSPSLHWSYGLRP 45
Db 2 HWSYGLRPGSGSQDWSY-----GLRPGSGSHWSYGLRP 35

RESULT 11
US-09-305-924-11
; Sequence 11, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 18.1965 Seconds
(without alignments)
133.345 Million cell updates/sec

Title: US-09-848-834A-17
Perfect score: 248
Sequence: 1 XHWSYGLRPGSSGSLKLLS.....HRLEGVGPJHWSYGLRPX 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 117 | 47.2 | 42 | 1 | US-08-446-692-20 |
| 2 | 117 | 47.2 | 42 | 2 | US-08-488-351A-20 |
| 3 | 116 | 46.8 | 27 | 1 | US-08-446-692-19 |
| 4 | 116 | 46.8 | 27 | 2 | US-08-488-351A-19 |
| 5 | 116 | 46.8 | 27 | 3 | US-09-100-414B-36 |
| 6 | 116 | 46.8 | 27 | 3 | US-09-303-323-36 |
| 7 | 116 | 46.8 | 27 | 4 | US-09-770-014-36 |
| 8 | 116 | 46.8 | 45 | 1 | US-08-446-692-33 |
| 9 | 116 | 46.8 | 45 | 2 | US-08-488-351A-33 |
| 10 | 113 | 45.6 | 27 | 3 | US-09-100-414B-43 |
| 11 | 113 | 45.6 | 27 | 3 | US-09-303-323-43 |
| 12 | 113 | 45.6 | 27 | 4 | US-09-770-014-43 |
| 13 | 110 | 44.4 | 31 | 3 | US-09-100-414B-55 |
| 14 | 110 | 44.4 | 31 | 3 | US-09-303-323-55 |
| 15 | 110 | 44.4 | 31 | 4 | US-09-770-014-55 |
| 16 | 108 | 43.5 | 27 | 3 | US-09-100-414B-41 |
| 17 | 108 | 43.5 | 27 | 3 | US-09-100-414B-47 |
| 18 | 108 | 43.5 | 27 | 3 | US-09-303-323-41 |
| 19 | 108 | 43.5 | 27 | 3 | US-09-303-323-47 |
| 20 | 108 | 43.5 | 27 | 4 | US-09-770-014-41 |
| 21 | 108 | 43.5 | 27 | 4 | US-09-770-014-47 |
| 22 | 108 | 43.5 | 45 | 3 | US-09-100-414B-45 |
| 23 | 108 | 43.5 | 45 | 3 | US-09-303-323-45 |
| 24 | 108 | 43.5 | 45 | 4 | US-09-770-014-45 |
| 25 | 107 | 43.1 | 31 | 3 | US-09-100-414B-59 |
| 26 | 107 | 43.1 | 31 | 3 | US-09-303-323-59 |
| 27 | 107 | 43.1 | 31 | 4 | US-09-770-014-59 |

| | | | | | | |
|----|-------|------|----|---|-------------------|-------------------|
| 28 | 106 | 42.7 | 28 | 3 | US-09-100-414B-38 | Sequence 38, Appl |
| 29 | 106 | 42.7 | 28 | 3 | US-09-303-323-38 | Sequence 38, Appl |
| 30 | 106 | 42.7 | 28 | 4 | US-09-770-014-38 | Sequence 38, Appl |
| 31 | 106 | 42.7 | 40 | 2 | US-08-460-502-10 | Sequence 10, Appl |
| 32 | 106 | 42.7 | 40 | 4 | US-09-148-711A-10 | Sequence 10, Appl |
| 33 | 105 | 42.3 | 31 | 3 | US-09-100-414B-53 | Sequence 53, Appl |
| 34 | 105 | 42.3 | 31 | 3 | US-09-303-323-53 | Sequence 53, Appl |
| 35 | 105 | 42.3 | 31 | 4 | US-09-770-014-53 | Sequence 53, Appl |
| 36 | 105 | 42.3 | 47 | 3 | US-09-100-414B-60 | Sequence 60, Appl |
| 37 | 105 | 42.3 | 47 | 3 | US-09-303-323-60 | Sequence 60, Appl |
| 38 | 105 | 42.3 | 47 | 4 | US-09-770-014-60 | Sequence 60, Appl |
| 39 | 105 | 42.3 | 49 | 3 | US-09-100-414B-57 | Sequence 57, Appl |
| 40 | 105 | 42.3 | 49 | 3 | US-09-303-323-57 | Sequence 57, Appl |
| 41 | 105 | 42.3 | 49 | 4 | US-09-770-014-57 | Sequence 57, Appl |
| 42 | 103.5 | 41.7 | 25 | 1 | US-08-446-692-17 | Sequence 17, Appl |
| 43 | 103.5 | 41.7 | 25 | 2 | US-08-488-351A-17 | Sequence 17, Appl |
| 44 | 103 | 41.5 | 27 | 3 | US-09-100-414B-50 | Sequence 50, Appl |
| 45 | 103 | 41.5 | 27 | 3 | US-09-303-323-50 | Sequence 50, Appl |

ALIGNMENTS

RESULT 1
US-08-446-692-20
; Sequence 20, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-446-692-20

Query March 47.2%; Score 117; DB 1; Length 42;
Best Local Similarity 82.8%; Pred. No. 9.5e-09;
Matches 24; Conservative 1; Mismatches 2; Indels 1;
Gaps 1;

QY 18 LLSIKGVIVHRLEGVGPJHWSYGLRP 46
Db 15 VLSIKGVIVHRLEGVGPJHWSYGLRP 41

RESULT 2

US-08-488-351A-20
 ; Sequence 20, Application US/08488351A
 ; Patent No. 5843446
 ; GENERAL INFORMATION:
 ; APPLICANT: Ladd, Anna
 ; APPLICANT: Wang, Chang Yi
 ; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION NUMBER: US/08/488,351A

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; APPLICATION NUMBER: US 08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION NUMBER: US/08/446,692

; FILING DATE: 7-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria C.H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4146 US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 415-8745

; TELEFAX: (516) 751-6849

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 27 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-446-692-19

Query Match 46.8%; Score 116; DB 1; Length 27;

Best Local Similarity 85.7%; Pred. No. 7.5e-09;

Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

US-08-446-692-19

QY 19 LSEIKGVIVHRLGVGSPSLHWSYGLRP 46

Db 1 LSEIKGVIVHRLGVGSGE--HWSYGLRP 26

RESULT 4

US-08-488-351A-19

; Sequence 19, Application US/08488351A

; Patent No. 5843446

; GENERAL INFORMATION:

; APPLICANT: Ladd, Anna

; APPLICANT: Wang, Chang Yi

; APPLICANT: Zamb, Timothy

; TITLE OF INVENTION: Immunogenic LHRH peptide constructs

; and synthetic universal immune stimulators for vaccines

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Maria C.H. Lin

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION NUMBER: US/08/488,351A

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA: US 08/446,692

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

FILING DATE: 7-JUN-1995

CLASSIFICATION: 424

APPLICATION NUMBER: US 08/446,692

Best Local Similarity 85.7%; Pred. No. 7.5e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46
Db 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 6
US-09-303-323-36
; Sequence 36, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-36

Query Match 46.8%; Score 116; DB 3; Length 27;
Best Local Similarity 85.7%; Pred. No. 7.5e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46
Db 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 7
US-09-770-014-36
; Sequence 36, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-19

Query Match 46.8%; Score 116; DB 2; Length 27;
Best Local Similarity 85.7%; Pred. No. 7.5e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 19 LSEIKGVIVHRLEGVGPGLHWSYGLRP 46
Db 1 LSEIKGVIVHRLEGVGPGLHWSYGLRP 26

RESULT 5
US-09-100-414B-36
; Sequence 36, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-36

Query Match 46.8%; Score 116; DB 3; Length 27;

```
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/770,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-770-014-36

Query Match 46.8%; Score 116; DB 4; Length 27;
Best Local Similarity 85.7%; Pred. No. 7.5e-09;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Oy 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
Db 1 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 26

RESULT 8
US-08-446-692-33
Sequence 33, Application US/08446692
Patent No. 5759551
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-33

Query Match 46.8%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.4e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Oy 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 46
Db 19 LSEIKGVIVHRLEGVGPSSLHWSYGLRP 44

RESULT 9
US-08-488-351A-33
Sequence 33, Application US/08488351A
Patent No. 5843446
GENERAL INFORMATION:
APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,351A
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/446,692
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/229,275
FILING DATE: 14-APR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/057,166
FILING DATE: 27-APR-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4146 US2
TELEPHONE: (212)415-8745
TELEFAX: (516)751-6849
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-488-351A-33

Query Match 46.8%; Score 116; DB 2; Length 45;
Best Local Similarity 85.7%; Pred. No. 1.4e-08;
Matches 24; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
```



```

: GENERAL INFORMATION:
: APPLICANT: Wang, Chang Yi
: TITLE OF INVENTION: NOVEL LHRH PEPTIDE
: TITLE OF INVENTION: IMMUNOGENS
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morgan & Finnegan, L.L.P.
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10154-0054
: COMPUTER READABLE FORM:

```

ZIP: 10154-0034
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/770,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-43

Query Match      45.6%; Score 113; DB 4; Length 27;
Best Local Similarity 82.1%; Pred. No. 1.9e-08;
Matches 23; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 19 LSEIKGVIVHRLGVGPGSLHWSYGLRP 46
Db 1 LSEIKGVIVHKLGVGGE--HWSYGLRP 26

RESULT 13
US-09-100-414B-55
; Sequence 55, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-55

Query Match      44.4%; Score 110; DB 3; Length 31;
Best Local Similarity 78.6%; Pred. No. 5.6e-08;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 19 LSEIKGVIVHRLGVGPGSLHWSYGLRP 46
Db 3 LSEIKGVIVHKLGVGGEHWSYGLRP 30

RESULT 14
US-09-303-323-55
; Sequence 55, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
```

```
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-55

Query Match      44.4%; Score 110; DB 4; Length 31;
Best Local Similarity 78.6%; Pred. No. 5.6e-08;
Matches 22; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      19 LSEIKGVIVHLEGVGEPGLHWSYGLRP 46
      |||||:|||||:|||||
Db      3 LSEIKGVIVHKLGVLFGEHWSYGLRP 30

Search completed: March 10, 2004, 09:28:56
Job time : 18.1965 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 15.5642 Seconds
(without alignments)
309.015 Million cell updates/sec

Title: US-09-848-834A-18
Perfect score: 274
Sequence: 1 XHWSYGLRPGSGPSLFNFF.....VSASHLEGPLSHWSYGLRXP 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 114 | 41.6 | 1315 | 1 BTCLTN | tentoxilysin (EC 3 |
| 2 | 67 | 24.5 | 1297 | 2 S39791 | neurotoxin - Clost |
| 3 | 65 | 23.7 | 1286 | 1 BTCLAB | bontoxilysin (EC 3 |
| 4 | 64 | 23.4 | 1268 | 2 S33411 | botulinum neurotox |
| 5 | 64 | 23.4 | 1291 | 2 I40631 | non-proteolytic bo |
| 6 | 63 | 23.0 | 369 | 2 S48109 | neurotoxin type F |
| 7 | 63 | 23.0 | 1274 | 2 I40813 | neurotoxin type F |
| 8 | 63 | 23.0 | 1291 | 1 A48940 | bontoxilysin (EC 3 |
| 9 | 62.5 | 22.8 | 1291 | 2 S46431 | botulinum neurotox |
| 10 | 62.5 | 22.8 | 1291 | 2 A49777 | botulinum neurotox |
| 11 | 62 | 22.6 | 1296 | 2 I40645 | botulinum neurotox |
| 12 | 61.5 | 22.4 | 91 | 2 JCR793 | medaka-type gonad |
| 13 | 61.5 | 22.4 | 92 | 1 RHRTG | gonadoliberin prec |
| 14 | 61 | 22.3 | 366 | 2 S48110 | neurotoxin type F |
| 15 | 61 | 22.3 | 502 | 2 T36589 | probable transmem |
| 16 | 59.5 | 21.7 | 1285 | 2 S70582 | botulinum neurotox |
| 17 | 59.5 | 21.7 | 1999 | 2 AB2018 | hypothetical prote |
| 18 | 59 | 21.5 | 67 | 2 I78541 | gonadoliberin prec |
| 19 | 58 | 21.5 | 92 | 1 RHVPG | gonadoliberin - pi |
| 20 | 58 | 21.2 | 10 | 1 RHVPG | gonadoliberin - pi |
| 21 | 58 | 21.2 | 10 | 1 RHVSG | gonadoliberin - sh |
| 22 | 58 | 21.2 | 89 | 2 I51423 | gonadoliberin prec |
| 23 | 58 | 21.2 | 90 | 1 RHVSG | gonadoliberin prec |
| 24 | 57.5 | 21.0 | 352 | 1 JVV824 | coat protein VP2 - |
| 25 | 57.5 | 21.0 | 836 | 2 JF0248 | ATP-binding cass |
| 26 | 57 | 20.8 | 665 | 2 H83403 | hypothetical prote |
| 27 | 56.5 | 20.6 | 98 | 2 I50739 | gonadotropin-relea |
| 28 | 56.5 | 20.6 | 345 | 2 A58519 | hypothetical 345 p |
| 29 | 56.5 | 20.6 | 367 | 2 S48106 | neurotoxin type E |

30 56.5 20.6 444 2 T24077
31 56.5 20.6 551 2 AC3572
32 56.5 20.6 658 2 S06744
33 56.5 20.6 1251 2 JH0256
34 56.5 20.6 1252 2 S21178
35 56 20.4 92 2 I50644
36 56 20.4 141 2 S48593
37 56 20.4 812 2 T01618
38 55.5 20.3 549 1 A47468
39 55.5 20.3 658 2 A89786
40 55.5 20.3 715 2 A41511
41 55 20.1 496 2 T38197
42 54.5 19.9 464 1 MNVWC
43 54.5 19.9 3122 2 T17202
44 54.5 19.9 5232 2 A45086
45 54 19.7 10 1 RHAQ1

ALIGNMENTS

RESULT 1

BTCLTN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N/Alternate names: tetanus neurotoxin
C/Species: Clostridium tetani
C/Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 03-Jun-2002
C/Accession: A25689; A25757; A25194; B25194; A60759; S69364
R/Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, EMBL J. 5, 2495-2502, 1986
A/Title: Tetanus toxin: Primary structure, expression in E. coli, and homology with b
A/Reference number: A25689; MUID:87053814; PMID:3536478
A/Accession: A25689
A/Molecule type: DNA
A/Residues: 1-1315 <BIS>
A/Cross-references: GB:X04436; NID:G40769; PIDN:CAA28033.1; PID:G40770
R/Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A/Title: The complete nucleotide sequence of tetanus toxin.
A/Reference number: A25757; MUID:87040747; PMID:3774547
A/Accession: A25757
A/Molecule type: DNA
A/Residues: 1-1315 <FAI>
A/Cross-references: GB:X06214; NID:G40773; PIDN:CAA29564.1; PID:G40774
A/Experimental source: strain CN3911
R/Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A/Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in
A/Reference number: A25194; MUID:86085672; PMID:3510187
A/Accession: A25194
A/Molecule type: DNA
A/Residues: 743-1315 <FA2>
A/Cross-references: GB:W12739; NID:G144920; PIDN:AAA23282.1; PID:G144921
A/Accession: B25194
A/Molecule type: protein
A/Residues: 865-894 <FA3>
R/Matsuda, M.; Lei, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A/Title: Isolation, purification, and characterization of fragment B, the NH-2-termina
A/Reference number: A60759; MUID:90035436; PMID:2478476
A/Accession: A60759
A/Molecule type: protein
R/Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A/Title: Delimitation of several DR-restricted tetanus toxin T cell epitopes.
A/Reference number: JS0098; MUID:89039918; PMID:2463305
A/Contents: annotation; epitope region
R/Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, I.
Nature 359, 832-835, 1992
A/Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteol
A/Reference number: S27125; MUID:93063293; PMID:1331807
A/Contents: annotation

hypothetical prote
lysine-CRNA ligase
staphylococcal
botulinum neurotox
botulinum neurotox
gonadoliberin I pr
probable methionyl
hypothetical prote
cytochrome-c oxida
staphylococcal
staphylococcal
probable myb-like
nonstructural prot
DNA-directed DNA p
HC-toxin synthetas
gonadoliberin I -

BTCLAB
Bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
N; Alternate names: Botulinum neurotoxin type A
C; Species: Clostridium botulinum
C; Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 18-Jun-1999
C; Accession: A35294; S09492; S68220; A33401; A53884; Ae0025; A27000
R; Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 255, 9153-9158, 1990
A; Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A; Reference number: A35294; MUID: 90264400; PMID: 2150960
A; Accession: A35294
A; Molecule type: DNA

| | | | |
|-----------------------|-------|----------------|-------------|
| Query Match | 25.7% | Score 0.5 | Length 1200 |
| Best Local Similarity | 56.2% | Pred. No. 7.4; | |
| Matches | 9; | Conservative | 5; |
| Mismatches | 2; | Indels | 0; |
| Gaps | 0; | | |

```
QY 15 SLFNNFTVSFWLRVPK 30
      |::|::|::|::|::|
Db 936 SMYENFTSFWIRIPK 951
      |::|::|::|::|::|

RESULT 4
botulinum neurotoxin type F - Clostridium barati
C:Species: Clostridium barati
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.
FEMS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
A:Reference number: S33411; MUID:93252228; PMID:8486245
A:Accession: S33411
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
A:Superfamily: tetanus toxin
C:Keywords: neurotoxin

      Query Match 23.4%; Score 64; DB 2; Length 1268;
      Best Local Similarity 62.5%; Pred. No. 9.7;
      Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      |::|::|::|::|::|
Db 920 SRYQNFSFWIRIPK 935
      |::|::|::|::|::|

RESULT 5
I40631
non-protective botulinum neurotoxin type B precursor - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40631; S48103; S48104; S36015
R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A:Title: Nucleotide sequence of the gene coding for non-protective Clostridium botulinum
A:Reference number: I40631; MUID:9412659; PMID:7764370
A:Accession: I40631
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1291 <RES>
A:Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific id
A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48103
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 634-761, 'E', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAM1>
A:Cross-references: EMBL:X70814; NID:g407778; PIDN:CAA50145.1; PID:g407779
A:Experimental source: non-protective strain 2129B (Scott)
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1993
A:Accession: S48104
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>
A:Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
A:Experimental source: non-protective strain Eklund 2B (Colworth 229)
C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic s
C:Genetics:
A:Gene: bont/b
C:Superfamily: tetanus toxin
C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LIGHT>
F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HY>
F:230,234/Binding site: zinc (His) #status predicted

F:231/Active site: Glu #status predicted

      Query Match 23.4%; Score 64; DB 2; Length 1291;
      Best Local Similarity 62.5%; Pred. No. 9.9;
      Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      |::|::|::|::|::|
Db 921 SMFLDFSFWIRIPK 936
      |::|::|::|::|::|

RESULT 6
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C>Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
R:Campbell, K.D.; Collins, M.D.; East, A.K.
C:Accession: S48109
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48109
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-369 <CAM>
A:Cross-references: EMBL:X70820; NID:g407790; PIDN:CAA50151.1; PID:g407791
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Superfamily: tetanus toxin

      Query Match 23.0%; Score 63; DB 2; Length 369;
      Best Local Similarity 56.2%; Pred. No. 3.2;
      Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      |::|::|::|::|::|
Db 295 SRYQNFSFWIRIPK 310
      |::|::|::|::|::|

RESULT 7
I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, D.
FEMS Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: I40644
A:Accession: I40813
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:Cross-references: GB:IM92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulin neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48108
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 634-1002 <CAM>
A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

      Query Match 23.0%; Score 63; DB 2; Length 1274;
      Best Local Similarity 56.2%; Pred. No. 13;
      Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
      |::|::|::|::|::|
Db 928 SRYQNFSFWIRIPK 943
      |::|::|::|::|::|
```

RESULT 8
A48940
Bontoxilysin (EC 3.4.24.69) B precursor - Clostridium botulinum
N/Alternate names: botulinum neurotoxin type B (BoNT/B)
C/Species: Clostridium botulinum
C/Date: 19-Dec-1993 #sequence revision 18-Nov-1994 #text_change 18-Jun-1999
A/Accession: A48940; S48105; S21575; S07155; S08562; S07128; S08574
R/Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A/Title: Molecular cloning of the Clostridium botulinum structural gene encoding the type B neurotoxin
A/Reference number: A48940; MUID:92384550; PMID:1514783
A/Accession: A48940
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1291 <WHB>
A/Cross-references: GB:M81186; NID:G144734; PIDN:AAA23211.1; PID:G144735
A/Experimental source: type B, Danish
A/Note: Sequence extracted from NCBI backbone (NCBI:112080, NCBI:112081); this publication is available in the EMBL database
R/Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A/Title: Gene probes for identification of the botulin neurotoxin gene and specific identification of the Clostridium botulinum type B and comparison of the Clostridium botulinum type B and Clostridium botulinum type A sequences
A/Reference number: S48103; MUID:94013372; PMID:8408542
A/Accession: S48105
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 634-994 <AM>
A/Cross-references: EMBL:X70817; NID:G407782; PIDN:CAA50148.1; PID:G407783
A/Experimental source: proteolytic type B, strain NCTC 7273
R/Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A/Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison of the Clostridium botulinum type B and Clostridium botulinum type A sequences
A/Reference number: S21575
A/Accession: S21575
A/Molecule type: DNA
A/Residues: 36-217, 'G', 219-224, 'S', 226-246 <SA>
A/Cross-references: EMBL:Z11934; NID:G40383; PIDN:CAA77991.1; PID:G40384
R/Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K.
J. Biol. Chem. 267, 14721-14729, 1992
A/Title: Minimal essential domains specifying toxicity of the light chains of tetanus toxin
A/Reference number: A42871; MUID:92340509; PMID:1634516
A/Accession: A42871
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-313, 'S', 315-451 <KUR>
A/Experimental source: strain Okra
A/Note: sequence extracted from NCBI backbone (NCBI:109365)
R/DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A/Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with type A
A/Reference number: S07155; MUID:89000987; PMID:3139097
A/Accession: S07155
A/Molecule type: protein
A/Residues: 2-29, 'M', 31-45 <DAS>
A/Accession: S08562
A/Molecule type: protein
A/Residues: 442-463, 'R', 465-467 <DA2>
R/Schmidt, J.J.; Sathyanarayanan, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A/Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A/Reference number: S07128; MUID:85197963; PMID:3888113
A/Accession: S07128
A/Status: preliminary
A/Molecule type: protein
A/Residues: 2-16 <SCH1>
A/Accession: S08573
A/Status: preliminary
A/Molecule type: protein
A/Residues: 2-17 <SCH2>
A/Accession: S08574
A/Status: preliminary
A/Molecule type: protein
A/Residues: 442-459 <SCH3>

R/Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, B.
Nature 359, 832-835, 1992
A/Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of syntaxin
A/Reference number: S27125; MUID:93063293; PMID:1331807
A/Contents: annotation
A/Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses
C/Genetics:
A/Accession: S08562
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1291 <WHB>
A/Cross-references: GB:M81186; NID:G144734; PIDN:AAA23211.1; PID:G144735
A/Experimental source: type B, Danish
A/Note: Sequence extracted from NCBI backbone (NCBI:112080, NCBI:112081); this publication is available in the EMBL database
R/Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A/Title: Gene probes for identification of the botulin neurotoxin gene and specific identification of the Clostridium botulinum type B and comparison of the Clostridium botulinum type B and Clostridium botulinum type A sequences
A/Reference number: S48103; MUID:94013372; PMID:8408542
A/Accession: S48105
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 634-994 <AM>
A/Cross-references: EMBL:X70817; NID:G407782; PIDN:CAA50148.1; PID:G407783
A/Experimental source: proteolytic type B, strain NCTC 7273
R/Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A/Description: Partial amino acid sequence of botulinum neurotoxin type B and comparison of the Clostridium botulinum type B and Clostridium botulinum type A sequences
A/Reference number: S21575
A/Accession: S21575
A/Molecule type: DNA
A/Residues: 36-217, 'G', 219-224, 'S', 226-246 <SA>
A/Cross-references: EMBL:Z11934; NID:G40383; PIDN:CAA77991.1; PID:G40384
R/Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K.
J. Biol. Chem. 267, 14721-14729, 1992
A/Title: Minimal essential domains specifying toxicity of the light chains of tetanus toxin
A/Reference number: A42871; MUID:92340509; PMID:1634516
A/Accession: A42871
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-313, 'S', 315-451 <KUR>
A/Experimental source: strain Okra
A/Note: sequence extracted from NCBI backbone (NCBI:109365)
R/DasGupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A/Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with type A
A/Reference number: S07155; MUID:89000987; PMID:3139097
A/Accession: S07155
A/Molecule type: protein
A/Residues: 2-29, 'M', 31-45 <DAS>
A/Accession: S08562
A/Molecule type: protein
A/Residues: 442-463, 'R', 465-467 <DA2>
R/Schmidt, J.J.; Sathyanarayanan, V.; DasGupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A/Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A/Reference number: S07128; MUID:85197963; PMID:3888113
A/Accession: S07128
A/Status: preliminary
A/Molecule type: protein
A/Residues: 2-16 <SCH1>
A/Accession: S08573
A/Status: preliminary
A/Molecule type: protein
A/Residues: 2-17 <SCH2>
A/Accession: S08574
A/Status: preliminary
A/Molecule type: protein
A/Residues: 442-459 <SCH3>

A;Residues: 1-84,'P',86-1291 <HAU>
A;Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906
R;Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi, K.
Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
A;Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxin
A;Reference number: A35396; MUID:91024998; PMID:2222445
A;Accession: A35396
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-669,'R',671-1291 <TS1>
R;Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
submitted to the EMBL Data Library, December 1991
A;Description: Nucleotide sequence of the gene for one of the components of hemagglutinin
A;Reference number: S22163
A;Accession: S22166
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1291 <TS2>
R;Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K.
Appl. Environ. Microbiol. 57, 1168-1172, 1991
A;Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and who
A;Reference number: A49777; MUID:91282468; PMID:2059039
A;Accession: A49777
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-607 <TS3>
A;Cross-references: GB:D90210
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 22.8%; Score 62.5; DB 2; Length 1291;
Best Local Similarity 25.4%; Pred. No. 15;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;
QY 5 YGLRGSQGP-----SLFNFTVSFWLRVFKVSAHLEGPSL----- 41
DB 907 FDFKLGSQSGEDRGKVIQVQENIVNSYFSFSFWLRINK-WVSNLPGLTIIDSVKNN 965
QY 42 -HWSYGL 47
DB 966 SGWSIGI 972

RESULT 11
I40645
botulinum neurotoxin type A - Clostridium botulinum
C;Species: Clostridium botulinum
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C;Accession: I40645
R;Willems, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A;Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A;Reference number: I40645; MUID:94143603; PMID:8310180
A;Accession: I40645
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1296 <RES>
A;Cross-references: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match 22.6%; Score 62; DB 2; Length 1296;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNFTVSFWLRVFK 30
DB 936 SWIENFSFWLRIK 951

RESULT 12
JC7393

medaka-type gonadotropin-releasing hormone precursor - Japanese medaka
C;Species: Oryzias latipes (Japanese medaka)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 20-Jan-2003
C;Accession: JC7393
R;Okubo, K.; Amano, M.; Yoshiura, Y.; Suetake, H.; Aida, K.
Biochem. Biophys. Res. Commun. 276, 299-303, 2000
A;Title: A novel form of gonadotropin-releasing hormone in the medaka, Oryzias latipes
A;Reference number: JC7393
A;Contents: Brain
A;Accession: JC7393
A;Molecule type: mRNA
A;Residues: 1-91 <OKU>
A;Cross-references: DDBJ:AB041333
C;Comment: This protein plays the roles as a hypophysiotropic factor, and a physiologi
C;Genetics:
A;Gene: mdgnrh
C;Superfamily: gonadoliberin
C;Keywords: brain

Query Match 22.4%; Score 61.5; DB 2; Length 91;
Best Local Similarity 32.7%; Pred. No. 0.99;
Matches 17; Conservative 7; Mismatches 23; Indels 5; Gaps 1;

QY 2 HWSYGLRPGSSGSLFNFTVSFWLRV-----PKVSASHLEGPSLHWSYGLR 48
DB 23 HWSFGUSPGKRELKYPNTLENQRLNSNTPCSDLSHLESLAKYIRIK 74

RESULT 13
RHRTG

N;Alternase names: gonadoliberin-associated protein (GAP); gonadotropin releasing horm
N;Contains: gonadoliberin; prolactin release-inhibiting factor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C;Accession: A40147; B26173; A48410
R;Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A;Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic
A;Reference number: A40147; MUID:8938461; PMID:2476669
A;Accession: A40147
A;Molecule type: DNA
A;Residues: 1-92 <BON>
A;Cross-references: GB:M31670; NID:g204447; PIDN:AAA41264.1; PID:g204448
R;Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A;Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonad
A;Reference number: A94090; MUID:86094338; PMID:2867548
A;Accession: B26173
A;Molecule type: mRNA
A;Residues: 1-92 <ADE>
A;Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446
R;Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A;Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormo
A;Reference number: A48410; MUID:93105480; PMID:1468115
A;Accession: A48410
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-92 <MAI>
A;Cross-references: GB:S50870; NID:g262059; PIDN:AA824572.1; PID:g262060
A;Experimental source: thymus
A;Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
C;Genetics:
A;Introns: 47/3; 79/3
C;Function:
A;Description: stimulates pituitary secretion of lutropin and follitropin
A;Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-92/Product: progadoliberin #status predicted <PGN>
F;24-33/Product: gonadoliberin #status predicted <GLN>

Job time : 16.623 secs

F;37-92/Product: prolactin release-inhibiting factor #status predicted <PIP>
F;24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F;33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 22.4%; Score 61.5; DB 1; Length 92;
Best Local Similarity 51.7%; Pred. No. 1;
Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;

QY 28 VPKVASA-----HLEG-PSLHWSYGLRP 49
: : : : :
Db 4 IPKMAAVLLTVCLGCSQHWSYGLRP 32

RESULT 14

S48110
neurotoxin type F - Clostridium botulinum (fragment)
C/Species: Clostridium botulinum
C/Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C/Accession: S48110
R/Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A/Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A/Reference number: S48103; MUID:94013372; PMID:8408542
A/Accession: S48110
A/Status: preliminary; translation not shown
A/Molecule type: DNA
A/Residues: 1-366 <CAM>
A/Cross-references: EMBL:X70821; NID:g407792; PIDN:CAA50152.1; PID:g407793
C/Superfamily: tetanus toxin
C/Keywords: neurotoxin

Query Match 22.3%; Score 61; DB 2; Length 366;
Best Local Similarity 57.1%; Pred. No. 5.6;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 17 FNNFTVSFWLRVPK 30
: : : : :
Db 297 YQNFISFWVRPK 310

RESULT 15

T36589
probable transmembrane protein - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C/Accession: T36589
R/Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A/Reference number: Z21575
A/Accession: T36589
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-502 <OLI>
A/Cross-references: EMBL:AJ049826; PIDN:CAB42730.1; GSPDB:GNO0070; SCOEDB:SCH24.24
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SCH24.24
C/Superfamily: Mycobacterium tuberculosis probable membrane protein Rv0051

Query Match 22.3%; Score 61; DB 2; Length 502;
Best Local Similarity 24.1%; Pred. No. 8;
Matches 19; Conservative 7; Mismatches 21; Indels 32; Gaps 3;

QY 2 HWSYGL-----RFGSSGPS-----LFNNFTVSFWLRVPK 30
: : : : :
Db 421 HWAIGLHLGLTFLCAMVVRDILMPDRDPVRRRTGDDDFSGVLDGAEDFVLGPAARTER 480

QY 31 VSASHLEGPSLHWSYGLRP 49
: : : : :
Db 481 PTA-HFEGPQVHWGKGQRP 498

Search completed: March 10, 2004, 09:16:52

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:53 ; Search time 9.14397 Seconds
(without alignments)
284.724 Million cell updates/sec

Title: US-09-848-834A-18

Perfect score: 274

Sequence: 1 XHWSYGLRPGSGPSLFNFF.....VSASHLEGPSLHWSYGLRXP 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 114 | 41.6 | 1314 | 1 | TETX_CLOTE |
| 2 | 67 | 24.5 | 1296 | 1 | BXG_CLOBO |
| 3 | 65 | 23.7 | 1295 | 1 | BXA1_CLOBO |
| 4 | 63 | 23.0 | 431 | 1 | PURA_LEGPN |
| 5 | 63 | 23.0 | 1274 | 1 | BXF_CLOBO |
| 6 | 63 | 23.0 | 1290 | 1 | BXB_CLOBO |
| 7 | 62.5 | 22.8 | 1290 | 1 | BXC1_CLOBO |
| 8 | 62 | 22.6 | 1051 | 1 | V21_AHSV6 |
| 9 | 62 | 22.6 | 1295 | 1 | BXA2_CLOBO |
| 10 | 61.5 | 22.4 | 91 | 1 | GON1_ORVLA |
| 11 | 61.5 | 22.4 | 92 | 1 | GON1_RAT |
| 12 | 60 | 21.9 | 92 | 1 | GON1_TUPGB |
| 13 | 59 | 21.5 | 67 | 1 | GON1_MACWU |
| 14 | 59 | 21.5 | 92 | 1 | GON1_HUMAN |
| 15 | 58.5 | 21.4 | 1250 | 1 | BXE_CLOBO |
| 16 | 58 | 21.2 | 61 | 1 | GON1_SREP |
| 17 | 58 | 21.2 | 63 | 1 | GON1_MESAU |
| 18 | 58 | 21.2 | 89 | 1 | GON1_XENLA |
| 19 | 58 | 21.2 | 90 | 1 | GON1_MOUSE |
| 20 | 58 | 21.2 | 90 | 1 | GON1_RANCA |
| 21 | 58 | 21.2 | 91 | 1 | GON1_PIG |
| 22 | 57.5 | 21.0 | 352 | 1 | COA2_SV40 |
| 23 | 57.5 | 21.0 | 842 | 1 | ASC6_HUMAN |
| 24 | 56.5 | 20.6 | 94 | 1 | GON1_HAPBU |
| 25 | 56.5 | 20.6 | 551 | 1 | SVK_BRUME |
| 26 | 56.5 | 20.6 | 551 | 1 | SVK_BRUSU |
| 27 | 56.5 | 20.6 | 658 | 1 | STC1_STAAU |
| 28 | 56.5 | 20.6 | 1250 | 1 | BXE_CLOBO |
| 29 | 56 | 20.4 | 92 | 1 | GON1_CHICK |
| 30 | 56 | 20.4 | 141 | 1 | AMPM_MYCCA |
| 31 | 56 | 20.4 | 469 | 1 | CG51_HUMAN |
| 32 | 55.5 | 20.3 | 549 | 1 | FIXN_BRAJA |
| 33 | 55.5 | 20.3 | 715 | 1 | STC2_STAAU |

34 54.5 19.9 464 1 VNSS_TSWV1
35 54.5 19.9 3122 1 DPOZ_MOUSE
36 54.5 19.9 5217 1 HTS1_COCCA
37 54 19.7 10 1 GON1_ALIMI
38 54 19.7 585 1 YH70_SYNY3
39 53.5 19.5 536 1 OPGD_XYLFT
40 53.5 19.5 654 1 PSTA_MYCGB
41 53.5 19.5 760 1 AMY_GLOAB
42 53 19.3 99 1 GON1_DICUA
43 53 19.3 449 1 VNSS_INSVN
44 53 19.3 746 1 PFEA_PSEAB
45 53 19.3 1196 1 BXCN_CLOBO

ALIGNMENTS

RESULT 1
TETX_CLOTE
ID TETX_CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylisin) [Contains:
DE Tetanus toxin light chain (Tetanus toxin chain L); Tetanus toxin heavy
DE chain (Tetanus toxin chain H)].
GN TETX OR CTP60.
OS Clostridium tetani.
OG Plasmid pE88, and Plasmid 75 Kbp.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
ON NCBI_TaxID=1513;
RX [1]
RP SEQUENCE FROM N.A.
RC PLASMID=75 Kbp.
RX MEDLINE=87053814; PubMed=336478;
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel M., Habermann E., Niemann H.;
RT "Tetanus toxin: primary structure, expression in E. coli, and
RT homology with botulinum toxins.";
RN EMBO J. 5:2495-2502(1986).
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911; PLASMID=75 Kbp;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RT "The complete nucleotide sequence of tetanus toxin.";
RN Nucleic Acids Res. 14:7809-7812(1986).
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88; PLASMID=pE88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease.";
RN Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
RP SEQUENCE OF 742-1314 FROM N.A.
RC PLASMID=75 Kbp;
RX MEDLINE=86085572; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RT "Cloning, nucleotide sequencing, and expression of tetanus toxin
RT fragment C in Escherichia coli.";
RN J. Bacteriol. 165:21-27(1986).
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=9021034; PubMed=2108021;
RA Krieglstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin.";

P26002 tomato spot
Q61493 mus musculu
Q01886 cochllobolu
P37041 alligator m
P73627 synechocyst
Q87927 xylella fas
P47651 mycoplasma
P23671 clostridium
Q91a10 dicentrarch
Q01811 impatiens n
Q05098 pseudomonas
P46081 clostridium

RL Bur. J. Biochem. 188:39-45(1990).
 RN [6]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=92037649; PubMed=1935979;
 RA Krieglstein K.G., Henchen A.H., Weller U., Habermann E.;
 RT "limited proteolysis of tetanus toxin. Relation to activity and
 identification of cleavage sites.";
 RN Bur. J. Biochem. 202:41-51(1991).
 RP [7]
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93010948; PubMed=1396558;
 RA Schiavo G., Poulain B., Rossetto O., Benfenati F., Tauc L.,
 RA Montecucco C.;
 RT "tetanus toxin is a zinc protein and its inhibition of
 RT neurotransmitter release and protease activity depend on zinc.";
 RN EMBO J. 11:3577-3583(1992).
 RP [8]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RN Nature 359:832-835(1992).
 RP [9]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
 RX MEDLINE=97475217; PubMed=9334741;
 RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
 RA Sax M.;
 RT "Structure of the receptor binding fragment HC of tetanus
 RT neurotoxin.";
 RN Nat. Struct. Biol. 4:788-792(1997).
 CC -1- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-|-PHE-77
 CC BOND OF SYNAPTOSOMAL-2.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 76-Gln-|-Phe-77 bond in
 CC synaptobrevin 2.
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -1- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
 CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
 CC AND ARE NON-TOXIC AFTER SEPARATION.
 CC -1- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
 CC GANGLIOSIDE RECEPTORS.
 CC -1- SIMILARITY: Belongs to peptidase family M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X04436; CAA28033.1; -;
 CC EMBL; X06214; CAA29564.1; -;
 CC EMBL; AF528097; AAC37454.1; -;
 CC EMBL; M12739; AAA23282.1; -;
 CC PIR; A25689; BTCLTN.
 CC PDB; 1AF9; 29-APR-98.
 CC PDB; 1A8D; 14-OCT-98.
 CC PDB; 1DOH; 27-MAR-00.
 CC PDB; 1DFQ; 24-MAR-00.
 CC PDB; 1DIW; 24-MAR-00.
 CC PDB; 1DLL; 24-MAR-00.
 CC PDB; 1FV3; 05-SEP-01.
 CC MEROPS; M27.001; -;
 CC InterPro; IPR008985; ConA like lec.gl.
 CC InterPro; IPR002160; Kunitz legume.
 CC InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR000395; Peptidase M27.
 DR Pfam; PF01742; Peptidase M27; 1.
 DR PRINTS; PR00760; BONTOKILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE_1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT MET 0
 FT CHAIN 1 456
 FT METAL 457 1314
 FT ACT SITE 232 233
 FT METAL 236 236
 FT TRANSMEM 226 246
 FT DISULFID 438 466
 FT DISULFID 1076 1092
 FT HELIX 876 882
 FT TURN 883 883
 FT STRAND 884 891
 FT TURN 892 893
 FT STRAND 894 897
 FT TURN 904 907
 FT TURN 909 910
 FT STRAND 912 915
 FT STRAND 920 925
 FT TURN 928 929
 FT STRAND 932 935
 FT HELIX 938 940
 FT TURN 941 946
 FT STRAND 949 956
 FT HELIX 962 968
 FT TURN 969 970
 FT STRAND 972 977
 FT STRAND 980 981
 FT HELIX 983 985
 FT STRAND 987 995
 FT TURN 996 997
 FT STRAND 998 1004
 FT TURN 1006 1007
 FT STRAND 1010 1016
 FT STRAND 1020 1020
 FT TURN 1021 1022
 FT STRAND 1031 1037
 FT TURN 1039 1040
 FT STRAND 1042 1047
 FT TURN 1048 1049
 FT STRAND 1050 1056
 FT TURN 1058 1059
 FT STRAND 1068 1074
 FT TURN 1079 1080
 FT STRAND 1082 1091
 FT HELIX 1097 1105
 FT TURN 1106 1107
 FT STRAND 1112 1112
 FT STRAND 1114 1114
 FT TURN 1116 1117
 FT STRAND 1120 1120
 FT TURN 1122 1122
 FT STRAND 1123 1124
 FT TURN 1127 1131
 FT HELIX 1132 1134
 FT TURN 1135 1136
 FT STRAND 1137 1141
 FT TURN 1144 1145
 FT STRAND 1148 1152
 FT TURN 1155 1158
 FT STRAND 1159 1162
 FT STRAND 1163 1166
 FT TURN 1173 1178
 FT STRAND 1184 1185
 FT TURN 1188 1188
 FT STRAND 1190 1190

DR InterPro; IPR000395; Peptidase M27.
 DR Pfam; PF01742; Peptidase M27; 1.
 DR PRINTS; PR00760; BONTOKILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE_1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT MET 0
 FT CHAIN 1 456
 FT METAL 457 1314
 FT ACT SITE 232 233
 FT METAL 236 236
 FT TRANSMEM 226 246
 FT DISULFID 438 466
 FT DISULFID 1076 1092
 FT HELIX 876 882
 FT TURN 883 883
 FT STRAND 884 891
 FT TURN 892 893
 FT STRAND 894 897
 FT TURN 904 907
 FT TURN 909 910
 FT STRAND 912 915
 FT STRAND 920 925
 FT TURN 928 929
 FT STRAND 932 935
 FT HELIX 938 940
 FT TURN 941 946
 FT STRAND 949 956
 FT HELIX 962 968
 FT TURN 969 970
 FT STRAND 972 977
 FT STRAND 980 981
 FT HELIX 983 985
 FT STRAND 987 995
 FT TURN 996 997
 FT STRAND 998 1004
 FT TURN 1006 1007
 FT STRAND 1010 1016
 FT STRAND 1020 1020
 FT TURN 1021 1022
 FT STRAND 1031 1037
 FT TURN 1039 1040
 FT STRAND 1042 1047
 FT TURN 1048 1049
 FT STRAND 1050 1056
 FT TURN 1058 1059
 FT STRAND 1068 1074
 FT TURN 1079 1080
 FT STRAND 1082 1091
 FT HELIX 1097 1105
 FT TURN 1106 1107
 FT STRAND 1112 1112
 FT STRAND 1114 1114
 FT TURN 1116 1117
 FT STRAND 1120 1120
 FT TURN 1122 1122
 FT STRAND 1123 1124
 FT TURN 1127 1131
 FT HELIX 1132 1134
 FT TURN 1135 1136
 FT STRAND 1137 1141
 FT TURN 1144 1145
 FT STRAND 1148 1152
 FT TURN 1155 1158
 FT STRAND 1159 1162
 FT STRAND 1163 1166
 FT TURN 1173 1178
 FT STRAND 1184 1185
 FT TURN 1188 1188
 FT STRAND 1190 1190

DR InterPro; IPR000395; Peptidase M27.
 DR Pfam; PF01742; Peptidase M27; 1.
 DR PRINTS; PR00760; BONTOKILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE_1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;
 KW 3D-structure; Complete proteome.
 FT INIT MET 0
 FT CHAIN 1 456
 FT METAL 457 1314
 FT ACT SITE 232 233
 FT METAL 236 236
 FT TRANSMEM 226 246
 FT DISULFID 438 466
 FT DISULFID 1076 1092
 FT HELIX 876 882
 FT TURN 883 883
 FT STRAND 884 891
 FT TURN 892 893
 FT STRAND 894 897
 FT TURN 904 907
 FT TURN 909 910
 FT STRAND 912 915
 FT STRAND 920 925
 FT TURN 928 929
 FT STRAND 932 935
 FT HELIX 938 940
 FT TURN 941 946
 FT STRAND 949 956
 FT HELIX 962 968
 FT TURN 969 970
 FT STRAND 972 977
 FT STRAND 980 981
 FT HELIX 983 985
 FT STRAND 987 995
 FT TURN 996 997
 FT STRAND 998 1004
 FT TURN 1006 1007
 FT STRAND 1010 1016
 FT STRAND 1020 1020
 FT TURN 1021 1022
 FT STRAND 1031 1037
 FT TURN 1039 1040
 FT STRAND 1042 1047
 FT TURN 1048 1049
 FT STRAND 1050 1056
 FT TURN 1058 1059
 FT STRAND 1068 1074
 FT TURN 1079 1080
 FT STRAND 1082 1091
 FT HELIX 1097 1105
 FT TURN 1106 1107
 FT STRAND 1112 1112
 FT STRAND 1114 1114
 FT TURN 1116 1117
 FT STRAND 1120 1120
 FT TURN 1122 1122
 FT STRAND 1123 1124
 FT TURN 1127 1131
 FT HELIX 1132 1134
 FT TURN 1135 1136
 FT STRAND 1137 1141
 FT TURN 1144 1145
 FT STRAND 1148 1152
 FT TURN 1155 1158
 FT STRAND 1159 1162
 FT STRAND 1163 1166
 FT TURN 1173 1178
 FT STRAND 1184 1185
 FT TURN 1188 1188
 FT STRAND 1190 1190

RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
RT components of Clostridium botulinum type A progenitor toxins.";
RL FEBS Lett. 376:41-44(1995).
RN [6]
RP SEQUENCE OF 1-16.
RX MEDLINE=94178501; PubMed=6370252;
RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequence of the heavy and light chains of
RT botulinum neurotoxin type A.";
RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN [7]
RP SEQUENCE OF 1-46.
RX Dasgupta B.R., Foley J., Niece R.;
RA "Partial sequence of the light chain of botulinum neurotoxin type A";
RT Biochemistry 26:4162-4162(1987).
RN [8]
RP SEQUENCE OF 1-5 AND 444-456.
RX MEDLINE=91120847; PubMed=2126206;
RA Dasgupta B.R., Dekleva M.L.;
RT "Botulinum neurotoxin type A: sequence of amino acids at the
RT N-terminus and around the nicking site.";
RL Biochimie 72:661-664(1990).
RN [9]
RP SEQUENCE OF 448-464 AND 872-895.
RX MEDLINE=89024662; PubMed=3178218;
RA Sartymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
RT halves and their partial sequences.";
RL Arch. Biochem. Biophys. 266:142-151(1988).
RN [10]
RP SEQUENCE OF 448-482.
RX MEDLINE=85285016; PubMed=3896784;
RA Shone C.C., Hambleton P., Melling J.;
RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
RT and purification of two tryptic fragments. Proteolytic action near
RT the COOH-terminus of the heavy subunit destroys toxin-binding
RT activity.";
RL Eur. J. Biochem. 151:75-82(1985).
RN [11]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94063091; PubMed=8243676;
RA Benfenati F., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT COOH-terminal peptide bonds.";
RL FEBS Lett. 335:99-103(1993).
RN [12]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94124495; PubMed=8294407;
RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
RN [13]
RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX MEDLINE=21556941; PubMed=11700044;
RA Rigoni M., Caccini P., Johnson E.A., Montecucco C., Rossetto O.;
RT "Site-directed mutagenesis identifies active-site residues of the
RT light chain of botulinum neurotoxin type A.";
RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX MEDLINE=98455071; PubMed=9783750;
RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT "Crystal structure of botulinum neurotoxin type A and implications
RT for toxicity.";
RL Nat. Struct. Biol. 5:898-902(1998).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle

CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release
CC results in flaccid paralysis, with frequent heart or respiratory
CC failure.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- COFACTOR: Binds 1 zinc ion per subunit.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for
CC the treatment of strabismus and blepharospasm associated with
CC dystonia and cervical dystonia. Also used for the treatment of
CC hemifacial spasm and a number of other neurological disorders
CC characterized by abnormal muscle contraction.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -1- SIMILARITY: Belongs to peptidase family M27.
CC -1- DATABASE: NAME=BOTOX product information Web site;
CC WWW=http://www.botox.com/index.jsp?hp=productinfo".
CC -1- DATABASE: NAME=protein Spotlight;
CC NOTE=Issue 19 of February 2002;
CC WWW=http://www.expasy.org/spotlight/articles/sptlt019.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X52066; CAA36289.1; -
DR EMBL; M30196; AAA23262.1; -
DR EMBL; X92973; CAA63551.1; -
DR EMBL; D67030; BAA11051.1; -
DR EMBL; M27892; AAA23269.1; -
DR PIR; A35294; BTCLAB.
DR PDB; 3BTA; 01-OCT-99.
DR MEROPS; M27.002; -
DR InterPro; IPR008985; ConA_like_lect_gl.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR006025; Pept_M_zn_BS.
DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOLILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW Pharmacological; 3D-structure.
FT INIT MET 0
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT METAL 222 223 ZINC (CATALYTIC).
FT ACT SITE 223 223
FT METAL 226 226 ZINC (CATALYTIC).
FT METAL 261 261 ZINC (CATALYTIC).
FT DISULFID 429 453 INTERCHAIN.
FT DISULFID 1234 1279
FT TRANSMEM 626 646
FT TRANSMEM 655 675
FT VARIANT 26 26
FT MUTAGEN 261 261
FT MUTAGEN 265 265
FT MUTAGEN 365 365
FT CONFLICT 1 1
FT CONFLICT 479 479
FT CONFLICT 875 875
FT CONFLICT 891 891
FT SEQUENCE 1295 AA; 149322 MW; 858342F54862579 CRC64;

Query Match 23.7%; Score 65; DB 1; Length 1295;
 Best Local Similarity 56.2%; Pred. No. 4.4;
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTSVFWRVPEK 30
 :::::|::|::|::|::|
 DB 935 SMYENFTSVWRVPEK 950

RESULT 4

ID PURA LEGPN STANDARD; PRT; 431 AA.
 AC ORNM2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP-aspartate ligase)
 DE (AdSS) (AMPase).
 GN PURA.
 OS Legionella pneumophila.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 OC Legionellaceae; Legionella.
 OX NCBI_TaxID=446;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rankin S., Li Z., Isberg R.R.;
 RT "Macrophage induced genes of Legionella pneumophila: protection from
 RT reactive intermediates and solute imbalance during intracellular
 RT growth";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Plays an important role in the de novo pathway of purine
 CC nucleotide biosynthesis.
 CC -!- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
 CC adenylosuccinate.
 CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -!- PATHWAY: AMP biosynthesis: first committed step.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; AF480918; AAM00648.1; -.
 DR HAMAP; MF 00011; -; 1.
 DR InterPro; IPR001114; Asucc_synthase.
 DR Pfam; PF00709; Adenylosucc_synth; 1.
 DR ProDom; PD001188; Asucc_synthase; 1.
 DR TIGRFAMs; TIGR00184; pura; 1.
 DR PROSITE; PS01266; ADENYLOSUCIN SYN 1; 1.
 DR PROSITE; PS00513; ADENYLOSUCIN SYN 2; 1.
 KW Purine biosynthesis; Ligase; GTP-binding; Metal-binding; Magnesium.
 FT NP_BIND 13 19
 FT ACT_SITE 141 141
 FT ACT_SITE 148 148
 FT METAL 14 14
 FT METAL 41 41
 SQ SEQUENCE 431 AA; 47381 MW; D446C19AED550774 CRC64;

Query Match 23.0%; Score 63; DB 1; Length 431;
 Best Local Similarity 36.4%; Pred. No. 2.3;
 Matches 12; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 17 FNNFTSVFWRVPEKVSASHLEGPSLHWSYGLRP 49
 :::::|::|::|::|::|
 DB 168 YHNFVLTQYFQPAVDLESLLGESLQWAEELRP 200

RESULT 5

ID BXF CLOBO STANDARD; PRT; 1274 AA.
 AC P30396;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BONT/F)
 DE (Bontoxilysin F).
 GN BOTF.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 23387;
 RX MEDLINE=93012902; PubMed=1398040;
 RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
 RA Roberts T.A., Thompson D.E.;
 RT "Sequence of the gene encoding type F neurotoxin of Clostridium
 RT botulinum";
 RL FEMS Microbiol. Lett. 75:225-230 (1992).
 RN [2]
 RP SEQUENCE OF 1-64 FROM N.A.
 RA STRAIN=Hobbs FT10;
 RX MEDLINE=94297488; PubMed=7764998;
 RA East A.K., Collins M.D.;
 RT "Conserved structure of genes encoding components of botulinum
 RT neurotoxin complex M and the sequence of the gene coding for the
 RT nontoxic component in nonproteolytic Clostridium botulinum type F";
 RL Curr. Microbiol. 29:69-77 (1994).
 RN [3]
 RP SEQUENCE OF 634-1002 FROM N.A.
 RA MEDLINE=94013372; PubMed=9408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulinum neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F";
 RL J. Clin. Microbiol. 31:2255-2262 (1993).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RA MEDLINE=94230352; PubMed=8175689;
 RA Yamasaki S., Baumeister A., Binz T., Biasi J., Link E., Cornille F.,
 RA Roques B., Fyfe E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RT F botulinum neurotoxins and tetanus toxin";
 RL J. Biol. Chem. 269:12764-12772 (1994).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-LYS-59
 CC BOND OF SYNAPTOSOMAL-1 AND -2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H). The light chain has the pharmacological activity,
 CC while the N- and C-terminal of the heavy chain mediate channel
 CC formation and toxin binding, respectively.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

-----
DR ENBL; M92906; AAA23263.1; -
DR ENBL; S73676; AAC60475.1; -
DR ENBL; X70820; CAA50151.1; -
DR ENBL; X70816; CAA50147.1; -
DR PIR; I40813; I40813.
DR PIR; S48109; S48109.
DR HSPP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR008985; ConA like lec.gl.
DR InterPro; IPR002180; Kunitz legume.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000395; Peptidase M27.
DR Pfam; PF01742; Peptidase M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT CHAIN 1 436 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
FT CHAIN 437 1274 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
FT METAL 227 227 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 228 228 BY SIMILARITY.
FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
FT D-SULFID 429 445 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 23.0%; Score 63; DB 1; Length 1274;
Best Local Similarity 56.2%; Pred. No. 7.7;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSFWLRVPK 30
Db 928 SRVQNFSSFWLRVPK 943

RESULT 6
ID BXB CLOBO STANDARD; PRT; 1290 AA.
AC P10844; P10843;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
DE (Bontoxilysin B).
GN BOTB.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;
RA Whelan S.M.; Elmore M.J.; Bodsworth N.J.; Brehm J.K.; Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence."
RL Appl. Environ. Microbiol. 58:2345-2354 (1992).
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN=NCTC 7273;
RA Szabo E.A.; Pemberton J.M.; Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN=NCTC 7273;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K.; East A.K.; Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."

```

```

RL J. Clin. Microbiol. 31:2255-2262 (1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
RA Dasgupta B.R.; Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin."
RL Biochimie 70:811-817 (1988).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.
RC STRAIN=OKRA;
RX MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J.; Sathyaamoorthy V.; Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E."
RL Arch. Biochem. Biophys. 238:544-548 (1985).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93054694; PubMed=1429690;
RA Schiavo G.; Rossetto O.; Santucci A.; Dasgupta B.R.; Montecucco C.;
RL J. Biol. Chem. 267:23479-23483 (1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G.; Benfenati F.; Poulain B.; Rossetto O.; de Laureto P.P.;
RL "Botulinum neurotoxins are zinc proteases."
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin."
RL Nature 359:832-835 (1992).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
CC SYNAPTOSOMAL-2.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H). The light chain has the pharmacological activity,
CC while the N- and C-terminal of the heavy chain mediate channel
CC formation and toxin binding, respectively.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -1- SIMILARITY: Belongs to peptidase family M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

EMBL; M81186; AAA23211.1; -
EMBL; Z11934; CAA77991.1; -
EMBL; X70817; CAA50148.1; -
PIR; A48940; A48940.
PDB; 1EPW; 01-NOV-00.
PDB; 1F31; 01-NOV-00.
PDB; 1F82; 16-AUG-00.
PDB; 1F83; 16-AUG-00.
PDB; 1FQ3; 08-DEC-00.
PDB; 1G9A; 13-NOV-02.
PDB; 1G9B; 13-NOV-02.
PDB; 1G9C; 13-NOV-02.
PDB; 1G9D; 13-NOV-02.
PDB; 111E; 21-NOV-01.

```

```

DR MEOPRS; M27.002; -.
DR InterPro; IPR008985; ConA_like lec_gl.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PRO0760; BONTOKILYSIN.
DR ProDom; PD001963; Bontokilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; transmembrane; Hydrolase; Metalloprotease; Zinc;
KW 3D-structure.
FT INIT_MET 0
FT CHAIN 1 440
FT CHAIN 441 1290
FT METAL 229
FT ACT_SITE 230 233
FT METAL 233 233
FT DISULFID 436 445
FT CONFLICT 29 217
FT CONFLICT 217 217
FT CONFLICT 224 224
FT CONFLICT 463 463
SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;

Query Match 23.0%; Score 63; DB 1; Length 1290;
Best Local Similarity 62.5%; Pred. No. 7.8;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNNFTVSWLRVPEK 30
DB 920 SVFLDFSVFWIRPK 935

RESULT 7
EX1_CLOBO STANDARD; PRT; 1290 AA.
AC P18640;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BoNT/C1)
DE (Bontokilysin C1).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90370487; PubMed=2204031;
RA Hauser D., Eklund M.W., Kurazona H., Binz T., Niemann H., Gill D.M.,
RA Boquet P., Popoff M.R.;
RT "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL Nucleic Acids Res. 18:4924-4924(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Type C Stockholm / C-ST;
RC MEDLINE=91024998; PubMed=222445;
RA Kimura K., Fujii N., Tazuki K., Murakami T., Indoh T.,
RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT "The complete nucleotide sequence of the gene coding for botulinum
type C1 toxin in the C-ST phage genome.";
RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
RN [3]
RP SEQUENCE OF 2-25.
RC STRAIN=Type C Stockholm / C-ST;
RX MEDLINE=68153072; PubMed=2450068;
RA Tezuka K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
RA Oguma K.;
RT "Establishment of a monoclonal antibody recognizing an antigenic site
common to Clostridium botulinum type B, C1, D, and E toxins and
tetanus toxin.";
RL Infect. Immun. 56:898-902(1988).
RN [4]

```

```

RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94038966; PubMed=7901002;
RT Blasi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.;
RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
cleaving HPC-1/syntaxin.";
RL EMBO J. 12:4821-4828(1993).
CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPETIDASE THAT CLEAVES SYNTAXIN.
CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
detected action on small molecule substrates.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
heavy chain (H). The light chain has the pharmacological activity,
while the N- and C-terminal of the heavy chain mediate channel
formation and toxin binding, respectively.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
botulinum neurotoxin: Types A, B, C1, D, E, F and G.
CC -!- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
BACTERIOPHAGE.
CC -!- SIMILARITY: Belongs to peptidase family M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X66433; CAA47060.1; -
CC EMBL; X72793; CAA51313.1; -
CC EMBL; X53751; CAA37780.1; -
CC EMBL; D90210; BAA14235.1; -
CC EMBL; X62389; CAA44263.1; -
CC HSP; P10845; 3BTA.
CC MEROPS; M27.002; -.
DR InterPro; IPR008985; ConA_like lec_gl.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PRO0760; BONTOKILYSIN.
DR ProDom; PD001963; Bontokilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 448
FT CHAIN 449 1290
FT METAL 228 228
FT ACT_SITE 229 229
FT METAL 232 232
FT DISULFID 436 452
FT CONFLICT 84 84
SQ SEQUENCE 1290 AA; 148734 MW; 71FBE379F7129E8 CRC64;

Query Match 22.8%; Score 62.5; DB 1; Length 1290;
Best Local Similarity 25.4%; Pred. No. 9;
Matches 17; Conservative 13; Mismatches 12; Indels 25; Gaps 3;

QY 5 YGLRGSSGSP-----SLFNNFTVSWLRVPEKVSASHLEGPSL-----41
DB 906 FDFKLGGSGEDRGKVIYTONENIVNMYESFSIFWIRNK-WVSNLPGYTIIDSVKNN 964
QY 42 -HWSYGL 47
DB 965 SCWSIGI 971

```



```

RESULT 8
ID_VP2_AHSV6 STANDARD; PRT; 1051 AA.
AC OVI024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer capsid protein VP2.
GN S2 OR L2.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
OS (serotype 6)).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=86060;
RN [1]
RP MEDLINE=98278311; PubMed=9617769;
RX Williams C.F., Inoue T., Lucas A.-M., Zanotto P., Roy P.;
RA "The complete sequence of four major structural proteins of African
RT horse sickness virus serotype 6: evolutionary relationships within
RT and between the orbiviruses."
RL Virus Res. 53:53-73(1998).
CC -1- FUNCTION: THE VP2 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP5)
CC WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID. IT IS THE
CC MAJOR TARGET OF THE HOST IMMUNOGENIC RESPONSE.
CC -1- SIMILARITY: Belongs to the reoviruses VP2 protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AF021235; AAC40994.1; -.
DR InterPro; IPR001742; Orbi_VP2.
DR Pfam; PFO0898; Orbi_VP2; 1.
KW Coat protein.
SQ SEQUENCE 1051 AA; 122326 MW; 2B04DB9B389F4B5F CRC64;

Query Match 22.6%; Score 62; DB 1; Length 1051;
Best Local Similarity 40.7%; Pred. No. 8.2;
Matches 11; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 11 SSGPSLFNNFTVFWLRVVKVASHLE 37
Db 630 TEGVTYFSKRFVSYWRVEKITKHL 656

RESULT 9
ID_BXA2_CLOBO STANDARD; PRT; 1295 AA.
AC Q45894; P77780;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BoNT/A)
DE (Bontoxilin A) (BOROX) [Contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=94143603; PubMed=8310180;
RA Williams A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
RT botulinum type A associated with infant botulism: comparison with

```

```

other clostridial neurotoxins.";
Res. Microbiol. 144:547-556(1993).
[2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=Kyoto-F;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC the C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-Gln-|-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release
CC results in flaccid paralysis, with frequent heart or respiratory
CC failure (By similarity).
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H) (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -1- SIMILARITY: Belongs to peptidase family M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X73423; CAA51824.1; -.
DR EMBL; X87974; CAA61234.1; -.
DR PIR; I40645; I40645.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR008985; ConA_like lec_gl.
DR InterPro; IPR002160; Kunitz_legume.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PFO1742; Peptidase_M27.
DR PRINTS; PR00760; BONTXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC PROTEASE; FALSE NEG.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT MET 0 0
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT METAL 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT ACT SITE 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 453 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1234 1279 BY SIMILARITY.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;

Query Match 22.6%; Score 62; DB 1; Length 1295;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 15 SLFNNFTVSWLRVVK 30
::: |||:::

```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; S50870; AAB24572.1; -
CC EMBL; M12579; AAA41263.1; -
CC EMBL; M1670; AAA41264.1; -
CC EMBL; M15527; AAA42141.1; ALT_SEQ.
CC EMBL; M15529; AAA42139.1; -
CC EMBL; M15528; -; NOT_ANNOTATED_CDS.
CC PIR; A40147; RHRTG.
CC InterPro; IPR002012; GnRH.
CC InterPro; IPR004079; GonadoliberinI.
CC Pfam; PF00446; GnRH; 1
CC PRINTS; PRO1541; GONADOLIBERNI.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;
Query Match 22.4%; Score 61.5; DB 1; Length 92;
Best Local Similarity 51.7%; Pred. No. 0.62;
Matches 15; Conservative 3; Mismatches 4; Indels 7; Gaps 2;

QY 28 VPKVSAS-----HLRG-PSLHWSYGLRP 49
DB 4 IPKLMAVLLTVLCGSSQHSYGLRP 32

RESULT 12

GON1_TUPGB STANDARD; PRT; 92 AA.
ID GON1_TUPGB
AC Q95375;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide II].
GN GNRH1 OR GnRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
OX NCBI_TaxID=37347;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGnRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GnRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; U63326; AAB16837.1; -
CC InterPro; IPR002012; GnRH.
CC InterPro; IPR004079; GonadoliberinI.
CC Pfam; PF00446; GnRH; 1
CC PRINTS; PRO1541; GONADOLIBERNI.
CC PROSITE; PS00473; GnRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal; Pyroglutamate carboxylic acid.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;
Query Match 21.9%; Score 60; DB 1; Length 92;
Best Local Similarity 52.4%; Pred. No. 0.95;
Matches 11; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 2 HWSYGLRPGS--SGPSLFNNF 20
DB 25 HWSYGLRPGGRNAENLDSF 45

RESULT 13

GON1_MACMU STANDARD; PRT; 67 AA.
ID GON1_MACMU
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor [Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I); GnRH-associated peptide I]
DE (Fragment).
GN GNRH1 OR GnRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=95124501; PubMed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
RT factor alpha and its receptor in the hypothalamus of female thes
RT macaques.";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
CC the secretion of both luteinizing and follicle-stimulating
CC hormones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the GnRH family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; S75918; AAB33096.1; -

```

DR PIR, I78541; I78541.
DR InterPro; IPR002012; GnRH.
DR InterPro; IPR004079; GonadoliberinI.
DR Pfam; PF00446; GnRH; 1.
DR PRINTS; PR01541; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
FT NON_TER 1 1
FT SIGNAL <1 5
FT CHAIN 6 >67
FT PEPTIDE 6 15
FT PEPTIDE 19 >67
FT ACT_SITE 8
FT MOD_RES 6 6
FT MOD_RES 15 15
FT NON_TER 67
FT SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;
Query Match 21.5%; Score 59; DB 1; Length 67;
Best Local Similarity 52.4%; Pred. No. 0.88; 5; Indels 2; Gaps 1;
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;
QY 2 HWSYGLRPGS--SGPSLFNF 20
DB 7 HWSYGLRPGKDAENLMSF 27

RESULT 14
GONI_HUMAN
ID GONI_HUMAN STANDARD; PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Progonadoliberin I precursor (Contains: Gonadoliberin I (LH-RH I)
DE (Luteinizing hormone-releasing hormone I) (Gonadotropin-releasing
DE hormone I) (GnRH I) (Luliberin I) (Gonadorelin); GnRH-associated
DE peptide I).
DE GnRH1 OR GnRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89366682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene."
RL Nucleic Acids Res. 17:6403-6403 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat."
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183 (1986).
RN [3]
RP SEQUENCE FROM N.A. AND VARIANT SER-16.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone."
RL Nature 311:666-668 (1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE=83128573; PubMed=6760865;
RA Tan L., Rousseau P.;

```

"The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta.";
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 [5]
 RN VARIANT SER-16.
 RX MEDLINE=99318093; PubMed=10391209;
 RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 Lander E.S.;
 RA "Characterization of single-nucleotide polymorphisms in coding regions
 of human genes.";
 RL Nat. Genet. 22:231-238 (1999).
 [6]
 RN ERRATUM.
 RP Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
 Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nemesh J., Ziaugra L.,
 Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
 Lander E.S.;
 RA Nat. Genet. 23:373-373 (1999).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 the secretion of both luteinizing and follicle-stimulating
 hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PHARMACEUTICAL: Available under the names Factrel (Ayerst Labs),
 Lutrepulse or Lutrelief (Perring Pharmaceuticals) and Relisorm
 (Serono).
 CC -!- SIMILARITY: Belongs to the GnRH family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see http://www.isb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 CC
 DR EMBL; X01059; CAA25526.1; -;
 DR EMBL; M12578; AAA35916.1; -;
 DR EMBL; X15215; CAA33285.1; -;
 DR PIR; S05308; RHUG.
 DR Genew; HGNC:4419; GNRH1.
 DR MIM; 152760.
 DR GO; GO:0005625; C:soluble fraction; TAS.
 DR GO; GO:0005183; F:lutinizing hormone-releasing factor activity; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007275; P:development; TAS.
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002012; GnRH.
 DR InterPro; IPR004079; GonadoliberinI.
 DR Pfam; PF00446; GnRH; 1.
 DR PRINTS; PR01541; GONADOLIBERNI.
 DR PROSITE; PS00473; GnRH; 1.
 DR Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pharmaceutical; Signal; Polymorphism;
 KW Pyroglutamate carboxylic acid.
 FT SIGNAL 1 23
 FT CHAIN 24 92
 FT PEPTIDE 24 33
 FT PEPTIDE 37 92
 FT ACT_SITE 26 26
 FT MOD_RES 24 24
 FT MOD_RES 33 33
 FT VARIANT 16 16
 FT SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
 SQ
 Query Match 21.5%; Score 59; DB 1; Length 92;
 Best Local Similarity 52.4%; Pred. No. 1.3;
 Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Qy 2 HWSYGLRPGS--SQPSLENNF 20
 Db 25 HWSYGLRPGGRKDAENLIDSF 45

RESULT 15
 BXE_CLOB0
 ID_BXE_CLOB0 STANDARD; PRT; 1250 AA.
 AC Q00496;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BoNT/E)
 DE (Bontoxilysin E).
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beluga;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulin neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92174922; PubMed=1541280;
 RA Whelan S.N., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 RT "The complete amino acid sequence of the Clostridium botulinum type-E
 RT neurotoxin, derived by nucleotide-sequence analysis of the encoding
 RT gene.";
 RL Eur. J. Biochem. 204:657-667(1992).
 RN [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz T., Kuzazono H., Willie M., Prevett J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [4]
 RP SEQUENCE OF 1-13.
 RX MEDLINE=85197963; PubMed=3898113;
 RA Schmidt J.J., Sathamoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [5]
 RP SEQUENCE OF 419-426.
 RX MEDLINE=90344918; PubMed=2116911;
 RA Gimenez J.A., Dasgupta B.R.;
 RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
 RT reveals the site trypsin nicks and homology with tetanus
 RT neurotoxin.";
 RL Biochimie 72:213-217(1990).
 RN [6]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8294407;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds";
 RL FEBS Lett. 335:99-103(1993).
 RN [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Binz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulin neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER

CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-|-ILE-
 CC 181 BOND IN SNAP-25.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H). The light chain has the pharmacological activity,
 CC while the N- and C-terminal of the heavy chain mediate channel
 CC formation and toxin binding, respectively.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -!- SIMILARITY: Belongs to peptidase family M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X62089; CAA43999.1; -;
 CC EMBL; X62683; CAA44558.1; -;
 CC PIR; S08575; S08575.
 CC FIR; S21178; S21178.
 CC HSP; P10845; 3BTA.
 CC MEROPS; M27.002; -;
 CC InterPro; IPR008985; ConA like lec.gl.
 CC InterPro; IPR002160; Kunitz legume.
 CC InterPro; IPR006025; Pept_M2n_BS.
 CC InterPro; IPR000395; Peptidase_M27.
 CC Pfam; PF01742; Peptidase_M27; 1.
 CC PRINTS; PR00760; BONTOLYLISN.
 CC ProDom; PD001963; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT MET 0
 FT CHAIN 1 421
 FT CHAIN 422 1250
 FT METAL 211 211
 FT ACT_SITE 212 212
 FT METAL 215 215
 FT DISULFID 411 425
 FT CONFLICT 176 176
 FT CONFLICT 197 197
 FT CONFLICT 339 339
 FT CONFLICT 772 772
 FT CONFLICT 962 963
 FT CONFLICT 966 966
 FT CONFLICT 1194 1194
 FT CONFLICT 1250 AA; 143712 NM; D9FCE26DDA041EB4 CRC64;
 SQ SEQUENCE 1250 AA; 143712 NM; D9FCE26DDA041EB4 CRC64;
 Query Match 21.4%; Score 58.5; DB 1; Length 1250;
 Best Local Similarity 23.6%; Pred. No. 27;
 Matches 13; Conservative 9; Mismatches 4; Indels 29; Gaps 2;
 Qy 17 FNNFTVSFWLRVP-----KVSASHLEGPSSLHWSY 45
 Db 911 YKNFSISFWLRIPNYDNKVNVEYTIINCRDNNSGWKVSLAHNE---LIWTF 962

Search completed: March 10, 2004, 09:13:58
 Job time : 9.14397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:54 ; Search time 49.4163 Seconds
(without alignments)
319.245 Million cell updates/sec

Title: US-09-848-834a-18

Perfect score: 274

Sequence: 1 XEWSXGLRFGSSGSLFNFN.....VSASHLEGPLSHWSYGLRPX 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 25:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 114 | 41.6 | 451 | Q91A13 | Q91A13 clostridium |
| 2 | 114 | 41.6 | 1310 | Q93N27 | Q93N27 clostridium |
| 3 | 64 | 23.4 | 361 | Q45848 | Q45848 clostridium |
| 4 | 64 | 23.4 | 361 | Q45846 | Q45846 clostridium |
| 5 | 64 | 23.4 | 441 | Q9X708 | Q9X708 clostridium |
| 6 | 64 | 23.4 | 1268 | Q45851 | Q45851 clostridium |
| 7 | 64 | 23.4 | 1291 | Q9ZAJ8 | Q9ZAJ8 clostridium |
| 8 | 64 | 23.4 | 1291 | Q93G71 | Q93G71 clostridium |
| 9 | 64 | 23.4 | 1291 | Q933K0 | Q933K0 clostridium |
| 10 | 64 | 23.4 | 1291 | Q08077 | Q08077 clostridium |
| 11 | 64 | 23.4 | 1291 | Q08G96 | Q08G96 clostridium |
| 12 | 64 | 23.4 | 2082 | Q9ACR4 | Q9ACR4 streptomyc |
| 13 | 62.5 | 22.8 | 1291 | Q93HT3 | Q93HT3 clostridium |
| 14 | 61.5 | 22.4 | 200 | Q45YU6 | Q45YU6 homo sapien |
| 15 | 61 | 22.3 | 228 | Q9J2H3 | Q9J2H3 macaca mula |
| 16 | 61 | 22.3 | 502 | Q9X8T8 | Q9X8T8 streptomyc |

| | | | | | |
|----|------|------|------|----|--------|
| 17 | 61 | 22.3 | 1278 | 2 | Q57236 |
| 18 | 60.5 | 22.1 | 363 | 10 | Q9FF71 |
| 19 | 59.5 | 21.7 | 657 | 10 | Q9XE83 |
| 20 | 59.5 | 21.7 | 1285 | 2 | Q45967 |
| 21 | 59.5 | 21.7 | 1285 | 2 | Q9LBR1 |
| 22 | 59.5 | 21.7 | 1999 | 16 | Q8YWB9 |
| 23 | 59 | 21.5 | 940 | 10 | Q7XLI6 |
| 24 | 58 | 21.2 | 91 | 13 | Q9FRH0 |
| 25 | 57.5 | 21.0 | 234 | 12 | Q9W9A5 |
| 26 | 57.5 | 21.0 | 234 | 12 | Q92837 |
| 27 | 57.5 | 21.0 | 352 | 12 | Q98VM1 |
| 28 | 57.5 | 21.0 | 352 | 12 | Q9W9F7 |
| 29 | 57.5 | 21.0 | 352 | 12 | Q80FH1 |
| 30 | 57.5 | 21.0 | 352 | 12 | Q91OV5 |
| 31 | 57.5 | 21.0 | 423 | 2 | Q7WSQ9 |
| 32 | 57.5 | 21.0 | 426 | 2 | Q8KN97 |
| 33 | 57.5 | 21.0 | 766 | 4 | Q96ME8 |
| 34 | 57.5 | 21.0 | 836 | 11 | Q70595 |
| 35 | 57.5 | 21.0 | 842 | 11 | Q9DC29 |
| 36 | 57.5 | 21.0 | 896 | 4 | Q9HAQ7 |
| 37 | 57 | 20.8 | 370 | 10 | Q8W414 |
| 38 | 57 | 20.8 | 430 | 2 | Q9XAV1 |
| 39 | 57 | 20.8 | 685 | 16 | Q912G5 |
| 40 | 57 | 20.8 | 866 | 5 | Q9VF20 |
| 41 | 57 | 20.8 | 1280 | 2 | Q9ZAJ5 |
| 42 | 56.5 | 20.6 | 94 | 13 | Q8JFY3 |
| 43 | 56.5 | 20.6 | 98 | 13 | Q805A5 |
| 44 | 56.5 | 20.6 | 113 | 2 | Q9KW76 |
| 45 | 56.5 | 20.6 | 140 | 2 | Q9K2Q0 |

ALIGNMENTS

RESULT 1
Q91A13
ID Q91A13 PRELIMINARY; PRT; 451 AA.
AC Q91A13, 2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Tetanus toxin (Fragment)
OS Clostridium tetani
OC Bacteria; Firmicutes; Clostridia; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154828; AAF73267.1; -
DR HSP; P04958; IABD.
DR GO; GO:0004866; F:Endopeptidase inhibitor activity; IEA.
DR InterPro; IPR008985; ConA like lec_g1.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR002160; Kunitz legume
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
FT NONTER
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F030E6CD8E CRC64;

Query Match 41.6%; Score 114; DB 2; Length 451;
Best Local Similarity 95.5%; Pred. No. 2.5e-06;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 LFNNFTVSVLWVRPKVSASHLE 37
Db 82 MFNNFTVSVLWVRPKVSASHLE 103

RESULT 2
Q93N27

```

SQ SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match      23.4%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred.No. 5.5;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNFTVSVFWLRVPK 30
DB 288 SMLDFSVSFWRIRPK 303
|.:|:|:|:|:|:|:|
|.:.|:|:|:|:|:|:|

RESULT 4
Q45846 PRELIMINARY; PRT; 361 AA.
ID AC Q45846;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Botulinum neurotoxin type B (Fragment).
DE BONT/B.
GN Clostridium botulinum.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
OX RN [1]
RP STRAIN=FROM N.A.
RC STRAIN=type B;
RX MEDLINE=34013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL; X70814; CAA50145.1; -.
DR HSPB; P10845; 3BTA.
DR GO; GO:0015070; F:coxin activity; IEA.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR Neurotoxin.
FT NON_TER 1
FT NON_TER 361
FT NON_TER 361
SQ SEQUENCE 361 AA; 42175 MW; 533EA98735CD98E1 CRC64;

Query Match      23.4%; Score 64; DB 2; Length 361;
Best Local Similarity 62.5%; Pred.No. 5.5;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 15 SLFNFTVSVFWLRVPK 30
DB 288 SMLDFSVSFWRIRPK 303
|.:|:|:~|:|:|:|
|.:.|:|:|:|:|:|:|

RESULT 5
Q9X708 PRELIMINARY; PRT; 441 AA.
ID AC Q9X708;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Botulinum neurotoxin type B (fragment).
DE BONT/B.
GN Clostridium botulinum.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
OX RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99343691; PubMed=10413679;
RX Lalli G., Herreros J., Osborne S.L., Montecucco C., Rossetto C.,
RX Schiavo G.;
RT "Functional characterisation of tetanus and botulinum neurotoxins
RT binding domains";
RL J. Cell Sci. 112:2715-2724(1999).
DR EMBL; AJ242628; CAB43706.1; -.
DR HSPB; P10845; 3BTA.
```

DT 01-MAY-1999 (TREMBlrel. 10, Created)

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Eklund 17B ATCC25765;
RX MEDLINE=94122659; PubMed=7764370;
RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT "Nucleotide sequence of the gene coding for non-proteolytic
Clostridium botulinum type B neurotoxin: comparison with other
Clostridial neurotoxins.";
RL Curr. Microbiol. 28:101-110(1994).
DR EMBL; X71343; CAA50482.1; -.
DR PIR; I40631; I40631.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27_002; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA_like lec_g1.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase_M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1291 AA; 150513 MW; 71B5CAFE23D69FAAA CRC64;
Query Match 23.4%; Score 64; DB 2; Length 1291;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 10; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
Oy 15 SLFNFTVTSFWLRVVK 30
|:|:|:|:|:|:|:
Db 921 SMELDFSVFWIRPK 936
RESULT 11
Q8GR96 PRELIMINARY; PRT; 1291 AA.
AC Q8GR96
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neurotoxin.
GN BONTB.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1491;
[1]
RN SEQUENCE FROM N.A.
RA Ihara H., Kohda T., Morimoto F., Tsukamoto K., Karasawa T.,
RA Nakamura S., Mukamoto M., Kozaki S.;
RT "Clostridium botulinum type B neurotoxin associated with infant
botulism.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB084152; BAC22064.1; -.
DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008985; ConA_like lec_g1.
DR InterPro; IPR002160; Kunitz legume.
DR InterPro; IPR000395; Peptidase_M27.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01742; Peptidase_M27; 1.
DR

IT'S NOT

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2004, 08:58:48 ; Search time 75.2918 Seconds
(without alignment)
187.635 Million cell updates/sec

Title: US-09-848-834A-18
Perfect score: 274
Sequence: 1 XHWSYGLRPGSGSPSLFNFP.....VSASHLGGPSLHNSYGLRXP 50

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 272 | 99.3 | 50 | AAU11429 | AAU11429 Synthetic |
| 2 | 199 | 72.6 | 37 | AAU11425 | AAU11425 Synthetic |
| 3 | 185 | 67.5 | 34 | AAU11421 | AAU11421 Synthetic |
| 4 | 152 | 55.5 | 32 | AAU1421 | AAU1421 Synthetic |
| 5 | 145 | 52.9 | 46 | AAU1430 | AAU1430 Synthetic |
| 6 | 141.5 | 51.6 | 47 | AAU1428 | AAU1428 Synthetic |
| 7 | 138.5 | 50.5 | 51 | AAU1431 | AAU1431 Synthetic |
| 8 | 126 | 46.0 | 708 | ABR82479 | ABR82479 Modified |
| 9 | 125 | 45.6 | 717 | ABR82478 | ABR82478 Modified |
| 10 | 123 | 44.9 | 109 | ABR20150 | ABR20150 Growth di |
| 11 | 123 | 44.9 | 194 | AAO30488 | AAO30488 Human TNF |
| 12 | 121.5 | 44.3 | 750 | AAU1425 | AAU1425 Synthetic |
| 13 | 121 | 44.2 | 216 | AAU1421 | AAU1421 Synthetic |
| 14 | 121 | 44.2 | 287 | AAU1428 | AAU1428 Synthetic |
| 15 | 121 | 44.2 | 537 | AAU1431 | AAU1431 Synthetic |
| 16 | 121 | 44.2 | 750 | AAU1425 | AAU1425 Synthetic |
| 17 | 120.5 | 44.0 | 109 | ABR20150 | ABR20150 Growth di |
| 18 | 120 | 43.8 | 188 | AAU1425 | AAU1425 Synthetic |
| 19 | 120 | 43.8 | 285 | AAU1421 | AAU1421 Synthetic |
| 20 | 119.5 | 43.6 | 109 | AAU1428 | AAU1428 Synthetic |
| 21 | 119.5 | 43.6 | 750 | AAU1425 | AAU1425 Synthetic |
| 22 | 119 | 43.4 | 122 | AAU1421 | AAU1421 Synthetic |
| 23 | 119 | 43.4 | 122 | AAU1428 | AAU1428 Synthetic |
| 24 | 118.5 | 43.2 | 31 | AAU1425 | AAU1425 Synthetic |
| 25 | 118.5 | 43.2 | 158 | AAU1421 | AAU1421 Synthetic |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 26 | 118.5 | 43.2 | 158 | 5 | ABO7282 | Human TNF |
| 27 | 118.5 | 43.2 | 693 | 3 | AAU1425 | AAU1425 Synthetic |
| 28 | 118.5 | 43.2 | 693 | 3 | AAU1421 | AAU1421 Synthetic |
| 29 | 118.5 | 43.2 | 750 | 3 | AAU1428 | AAU1428 Synthetic |
| 30 | 118.5 | 43.2 | 750 | 3 | AAU1431 | AAU1431 Synthetic |
| 31 | 118 | 43.1 | 158 | 5 | ABO7282 | Human TNF |
| 32 | 118 | 43.1 | 158 | 5 | ABO7279 | Human TNF |
| 33 | 118 | 43.1 | 703 | 3 | AAU1425 | AAU1425 Synthetic |
| 34 | 118 | 43.1 | 761 | 3 | AAU1421 | AAU1421 Synthetic |
| 35 | 117 | 42.7 | 124 | 3 | AAU1428 | AAU1428 Synthetic |
| 36 | 117 | 42.7 | 124 | 3 | AAU1431 | AAU1431 Synthetic |
| 37 | 117 | 42.7 | 160 | 4 | AAO30488 | AAO30488 Human TNF |
| 38 | 117 | 42.7 | 160 | 4 | AAO30495 | AAO30495 Human TNF |
| 39 | 117 | 42.7 | 750 | 3 | AAU1425 | AAU1425 Synthetic |
| 40 | 116 | 42.3 | 31 | 3 | AAU1425 | AAU1425 Synthetic |
| 41 | 116 | 42.3 | 43 | 4 | AAU1421 | AAU1421 Synthetic |
| 42 | 116 | 42.3 | 43 | 4 | AAU1428 | AAU1428 Synthetic |
| 43 | 116 | 42.3 | 72 | 4 | AAU1431 | AAU1431 Synthetic |
| 44 | 116 | 42.3 | 109 | 4 | AAU1425 | AAU1425 Synthetic |
| 45 | 116 | 42.3 | 136 | 4 | AAU1421 | AAU1421 Synthetic |

ALIGNMENTS

RESULT 1

AAU11429
ID AAU11429 standard; peptide; 50 AA.

XX AC AAU11429;

XX DT 12-MAR-2002 (first entry)

XX DE Synthetic immunogen peptide 10.

XX KW Gonadotrophin releasing hormone; GnRH; synthetic immunogen;

XX KW luteinising hormone releasing hormone; GnRH; synthetic immunogen;

XX KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;

XX KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;

XX KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.

XX OS Clostridium tetani.

XX OS Mammalia.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key

XX FT Peptide

XX FT Location/Qualifiers

XX FT 1..10

XX FT /note= "Gonadotrophin releasing hormone epitope (1..10 aa)"

XX FT Misc-difference 1

XX FT /label= OTHER

XX FT /note= "Other= Pyro-glutamic acid or 5-oxo proline"

XX FT Peptide

XX FT /note= "Spacer peptide"

XX FT Peptide

XX FT /note= "Tetanus toxoid (947-967 aa)"

XX FT Peptide

XX FT /note= "Spacer peptide"

XX FT Peptide

XX FT /note= "Gonadotrophin releasing hormone epitope (2-10 aa)"

XX FT Modified-site

XX FT /note= "Amidated glycine or glycylamide"

XX FT WO200185763-A2.

XX FT 15-NOV-2001.

XX FT 04-MAY-2001; 2001WO-US014363.

XX FT 05-MAY-2000; 2000US-0202328P.

XX FT

PA (APHT-) APHTON CORP.
 XX Grimes S, Michaeli D, Stevens VC;
 PI WPI; 2002-049440/06.
 XX
 XX Novel synthetic immunogen for inducing immune response against
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.
 XX
 XX Claim 11; Page 11; 43pp; English.
 PS The invention relates to a synthetic immunogen for inducing specific
 XX antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 XX Sequence 50 AA;
 SQ
 Query Match 99.3%; Score 272; DB 5; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1.6e-28;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPGSSGSLFNNFTVFWLVPKVSASHLEPSLHWSYGLRP 49
 DB 2 HWSYGLRPGSSGSLFNNFTVFWLVPKVSASHLEPSLHWSYGLRP 49
 RESULT 2
 AAU11425
 ID AAU11425 standard; peptide; 37 AA.
 XX
 AC AAU11425;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Synthetic immunogen peptide 6.
 XX
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX
 OS Clostridium tetani.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1. .10
 FT Misc-difference 1 /note= "Gonadotropin releasing hormone epitope"
 FT /label= OTHER
 FT Peptide 11. .16 /note= "Other= Pyro-glutamic acid or 5-oxo proline"
 FT /note= "Spacer peptide"
 FT Peptide 17. .37 /note= "Tetanus toxoid sequence (947-967 aa)"
 FT
 XX WO200185763-A2.
 PN 15-NOV-2001.
 XX

PF 04-MAY-2001; 2001WO-US014363.
 XX
 PR 05-MAY-2000; 2000US-0202328P.
 XX
 PA (APHT-) APHTON CORP.
 XX
 XX Grimes S, Michaeli D, Stevens VC;
 PI WPI; 2002-049440/06.
 XX
 XX Novel synthetic immunogen for inducing immune response against
 PT gonadotropin releasing hormone, comprises fusion peptide having
 PT promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
 PT or its analog.
 XX
 XX Claim 11; Page 9; 43pp; English.
 PS The invention relates to a synthetic immunogen for inducing specific
 XX antibodies against gonadotropin releasing hormone (GnRH) also known as
 CC luteinising hormone releasing hormone, LHRH) comprising a fusion peptide
 CC which comprises a promiscuous helper T-cell peptide epitope and
 CC immunomimic peptide epitope or its analogue. The synthetic immunogen is
 CC useful inducing an immune response against GnRH in an animal subject, and
 CC as such is useful as a contraceptive and in the treatment of diseases
 CC such as cancer (of the breast, uterus and other gynaecological cancer),
 CC endometriosis, uterine fibroids, benign prostatic hypertrophy and
 CC prostate cancer. The immunogen is effective in eliciting high and
 CC specific anti-GnRH antibody titres. The present sequence is a synthetic
 CC immunogen of the invention
 XX
 XX Sequence 37 AA;
 SQ
 Query Match 72.6%; Score 199; DB 5; Length 37;
 Best Local Similarity 100.0%; Pred. No. 5.4e-19;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 HWSYGLRPGSSGSLFNNFTVFWLVPKVSASHLE 37
 DB 2 HWSYGLRPGSSGSLFNNFTVFWLVPKVSASHLE 37
 RESULT 3
 AAU11421
 ID AAU11421 standard; peptide; 34 AA.
 XX
 AC AAU11421;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Synthetic immunogen peptide 2.
 XX
 KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
 KW luteinising hormone releasing hormone; LHRH; contraceptive;
 KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
 KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
 KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
 XX
 OS Clostridium tetani.
 OS Mammalia.
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1. .21 /note= "Tetanus toxoid sequence (947-967 aa)"
 FT Peptide 22. .25 /note= "Spacer peptide"
 FT Peptide 26. .34 /note= "Gonadotropin releasing hormone epitope"
 FT Modified-site 34 /note= "Amidated glycine or glycine amide"
 FT
 XX WO200185763-A2.
 PN

```

XX PD 15-NOV-2001.
XX PR
XX XX
XX PF 04-MAY-2001; 2001WO-US014363.
XX PA
XX PR 05-MAY-2000; 2000US-0202328P.
XX PA
XX XX (APHT-) APHTON CORP.
XX PI
XX GRIMES S, Michaeli D, Stevens VC;
XX WPI; 2002-049440/06.
XX DR
XX XX
XX XX Novel synthetic immunogen for inducing immune response against
XX Gonadotropin releasing hormone, comprises fusion peptide having
XX promiscuous helper T-cell peptide epitope and immunomimic peptide epitope
XX or its analog.
XX Claim 11; Page 7; 43pp; English.
XX CC The invention relates to a synthetic immunogen for inducing specific
XX antibodies against gonadotropin releasing hormone (GnRH) also known as
XX luteinising hormone releasing hormone (LHRH) comprising a fusion peptide
XX which comprises a promiscuous helper T-cell peptide epitope and
XX immunomimic peptide epitope or its analogue. The synthetic immunogen is
XX useful inducing an immune response against GnRH in an animal subject, and
XX as such is useful as a contraceptive and in the treatment of diseases
XX such as cancer (of the breast, uterus and other gynaecological cancer),
XX endometriosis, uterine fibroids, benign prostatic hypertrophy and
XX prostate cancer. The immunogen is effective in eliciting high and
XX specific anti-GnRH antibody titres. The present sequence is a synthetic
XX immunogen of the invention
XX SQ Sequence 34 AA;
XX
Query Match 67.5%; Score 185; DB 5; Length 34;
Best Local Similarity 100.0%; Pred. NO. 3.5e-17; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 0;
QY 17 FNNFTVSPFLRVKVSASHLEGPSSLHWSYGLRP 49
DB 1 FNNFTVSPFLRVKVSASHLEGPSSLHWSYGLRP 33
XX
RESULT 4
AAR62702
ID AAR62702 standard; peptide; 32 AA.
XX AC
XX AAR62702;
XX
DT 25-MAR-2003 (revised)
DT 10-SEP-1995 (first entry)
XX
DE LHRH-containing immunogenic peptide.
XX
KW Helper T cell epitope; universal immune stimulator; invasin; haptin;
KW vaccine; LHRH; luteinising hormone releasing hormone; prostate;
KW androgen-dependent carcinoma; antitumour; infertility; tetanus toxin.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..22
FT /note= "tetanus toxin helper T cell epitope"
FT Domain 23..32
FT /note= "LHRH haptin"
XX
FN WO9425060-A1.
XX
PD 10-NOV-1994.
XX
PF 28-APR-1994; 94WO-US004832.
XX

```

```

PR 27-APR-1993; 93US-00057166.
PR 14-APR-1994; 94US-00229275.
XX
XX (LADD/) LADD A E.
XX PA (WANG/) WANG C Y.
XX PR (ZAMB/) ZAMB T.
XX PI
XX LADD AE, Wang CY, Zamb T;
XX WPI; 1994-357910/44.
XX DR
XX XX Immunogenic luteinising hormone releasing hormone peptide(s) - that
XX suppress LHRH activity in males and females.
XX PT
XX XX Claim 8; Page 84; 213pp; English.
XX CC Synthetic immunogenic peptides are provided in which a universal immune
XX stimulator is linked to a peptide or protein haptin containing B cell
XX and/or cytotoxic T lymphocyte epitopes, giving a product which causes
XX potent immune responses to the coupled peptide or protein. The stimulator
XX consists of (A) a promiscuous helper T cell epitope (Th) which elicits an
XX immune response to the coupled peptide in members of a heterogeneous
XX population expressing diverse HLA phenotypes, and (B) an adjuvant peptide
XX sequence from the invasin protein of Yersinia. Spacer amino acid
XX sequences (e.g. Gly-Gly) can be provided between the invasin and Th
XX domains and between the immune stimulator and haptin components. When the
XX haptin is LHRH, then optionally the invasin domain can be omitted from
XX the immune stimulator component. The present sequence represents an LHRH-
XX containing, invasin-free immunogenic peptide as above which can be used
XX as a potent vaccine for treating e.g. prostatic hyperplasia, androgen-
XX dependent carcinoma, prostatic carcinoma, recurrent functional ovarian
XX endometriosis, benign uterine tumours, recurrent functional ovarian
XX cysts, (severe) premenstrual syndrome or oestrogen-dependent breast
XX cancer, or for induction of infertility. (Updated on 25-MAR-2003 to
XX correct FN field.)
XX SQ Sequence 32 AA;
XX
Query Match 55.5%; Score 152; DB 2; Length 32;
Best Local Similarity 87.9%; Pred. NO. 7.9e-13;
Matches 29; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 17 FNNFTVSPFLRVKVSASHLEGPSSLHWSYGLRP 49
DB 3 FNNFTVSPFLRVKVSASHLE----HWSYGLRP 31
XX
RESULT 5
AAR62702
ID AAR62702 standard; peptide; 46 AA.
XX AC
XX AAR62702;
XX
DT 12-MAR-2002 (first entry)
XX
DE Synthetic immunogen peptide 11.
XX
KW Gonadotropin releasing hormone; GnRH; synthetic immunogen;
KW luteinising hormone releasing hormone; LHRH; contraceptive;
KW promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
KW breast cancer; uterine cancer; gynaecological cancer; endometriosis;
KW uterine fibroid; benign prostatic hypertrophy; prostate cancer.
XX
OS Clostridium tetani.
XX OS Mammalia.
XX OS Synthetic.
XX OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 1..10
FT /note= "Gonadotropin releasing hormone epitope (1..10
FT aa)"
FT Misc-difference 1
XX

```

| | | |
|----------|--|---|
| KW | | uterine fibroid; benign prostatic hypertrophy; prostate cancer. |
| XX | | |
| OS | Plasmodium falciparum. | |
| OS | Mammalia. | |
| OS | Synthetic. | |
| OS | Chimeric. | |
| XX | | |
| PH | Key | Location/Qualifiers |
| FT | Peptide | 1..10 |
| FT | | /note= "Gonadotrophin releasing hormone epitope (1..10 aa)" |
| FT | Misc-difference 1 | |
| FT | | /label= OTHER |
| FT | | /note= "Other= Pyro-glutamic acid or 5-oxo proline" |
| FT | Peptide | 11..16 |
| FT | | /note= "Spacer peptide" |
| FT | Peptide | 17..34 |
| FT | | /note= "Malaria CSP protein (288-302 aa)" |
| FT | Peptide | 35..38 |
| FT | | /note= "Spacer peptide" |
| FT | Peptide | 39..47 |
| FT | | /note= "Gonadotrophin releasing hormone epitope (2-10 aa)" |
| FT | Modified-site | 47 |
| FT | | /note= "Amidated glycine or glycinamide" |
| XX | | |
| XX | WO200185763-A2. | |
| PX | | |
| PN | 15-NOV-2001. | |
| PD | | |
| XX | | |
| FF | 04-MAY-2001; 2001WO-US014363. | |
| PF | | |
| PR | 05-MAY-2000; 2000US-0202328P. | |
| XX | | |
| PA | (APHT-) APHTON CORP. | |
| XX | | |
| PI | Grimes S, Michaeli D, Stevens VC; | |
| XX | WPI; 2002-049440/06. | |
| XX | | |
| DR | Novel synthetic immunogen for inducing immune response against gonadotropin releasing hormone, comprises fusion peptide having promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analog. | |
| PT | | |
| PT | | |
| PT | | |
| PT | | |
| XX | | |
| PS | Claim 11; Page 11; 43pp; English. | |
| XX | | |
| CC | The invention relates to a synthetic immunogen for inducing specific antibodies against gonadotropin releasing hormone (GnRH) also known as luteinising hormone releasing hormone, LH(RH) comprising a fusion peptide which comprises a promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analogue. The synthetic immunogen is useful inducing an immune response against GnRH in an animal subject, and as such is useful as a contraceptive and in the treatment of diseases such as cancer (of the breast, uterus and other gynaecological cancer), endometriosis, uterine fibroids, benign prostatic hypertrophy and prostate cancer. The immunogen is effective in eliciting high and specific anti-GnRH antibody titres. The present sequence is a synthetic immunogen of the invention | |
| XX | | |
| QQ | Sequence 47 AA; | |
| | Query Match | 51.6%; Score 141.5; DB 5; Length 47; |
| | Best Local Similarity | 58.8%; Pred. No. 3.le-11; |
| | Matches | 30; Conservative 1; Mismatches 11; Indels 9; Gaps 2 |
| QY | 2 HWSYGLRPGCGSPSLFNNFTVSFWLRYPKYSASHL---EGPSLHWSYGLRP 49 | |
| | : | |
| Dd | 2 HWSYGLRPGCGSPSL-----KLISEIKGVIVHRLEGEVGPSLHWSYGLRP 46 | |
| | : | |
| RESULT 7 | | |
| AU011431 | | |

AAU11431 standard; peptide; 51 AA.
AAU11431;
12-MAR-2002 (first entry)
Synthetic immunogen peptide 12.
Gonadotrophin releasing hormone; GnRH; synthetic immunogen;
luteinising hormone releasing hormone; LHRH; contraceptive;
promiscuous helper T-cell peptide epitope; immunomimic peptide epitope;
breast cancer; uterine cancer; gynaecological cancer; endometriosis;
uterine fibroid; benign prostatic hypertrophy; prostate cancer.
Plasmodium falciparum.
Mammalia.
Synthetic.
Chimeric.
Key Location/Qualifiers
Peptide 1..10
/note= "Gonadotrophin releasing hormone epitope (1..10 aa)"
Misc-difference 1
/label= OTHER
/note= "Other= Pyro-glutamic acid or 5-oxo proline"
Peptide 11..16
/note= "Spacer peptide"
Peptide 17..36
/note= "Malaria CSP protein (378-398 aa)"
Peptide 37..42
/note= "Spacer peptide"
Peptide 43..51
/note= "Gonadotrophin releasing hormone epitope (2-10 aa)"
Modified-site 51
/note= "Amidated glycine or glycineamide"
WO200185763-A2.
15-NOV-2001.
04-MAY-2001; 2001WO-US014363.
05-MAY-2000; 2000US-0202328P.
(APHT-) APHTON CORP.
Grimes S, Michaeli D, Stevens VC;
WPI; 2002-049440/06.
Novel synthetic immunogen for inducing immune response against gonadotrophin releasing hormone, comprises fusion peptide having promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analog.
Claim 11; Page 12-13; 43pp; English.
The invention relates to a synthetic immunogen for inducing specific antibodies against gonadotrophin releasing hormone (GnRH) also known as luteinising hormone releasing hormone, LHRH) comprising a fusion peptide which comprises a promiscuous helper T-cell peptide epitope and immunomimic peptide epitope or its analog. The synthetic immunogen is useful inducing an immune response against GnRH in an animal subject, and as such is useful as a contraceptive and in the treatment of diseases such as cancer (of the breast, uterus and other gynaecological cancer), endometriosis, uterine fibroids, benign prostatic hypertrophy and prostate cancer. The immunogen is effective in eliciting high and specific anti-GnRH antibody titres. The present sequence is a synthetic immunogen of the invention
Sequence 51 AA;
QY 11 SSGPS-----LFNNFTVSFWLRVPKVSASHLEGPSLH 42
Query Match 46.0%; Score 136; DB 7; Length 708;
Best Local Similarity 58.4%; Pred. No. 8.4e-08;
Matches 26; Conservative 2; Mismatches 4; Indels 6; Gaps 1;
QY 11 SSGPS-----LFNNFTVSFWLRVPKVSASHLEGPSLH 42
Query Match 50.5%; Score 138.5; DB 5; Length 51;
Best Local Similarity 54.7%; Pred. No. 8.4e-11;
Matches 29; Conservative 4; Mismatches 11; Indels 9; Gaps 2;
QY 2 HWSYGLRPGSGPSLFNNFTVSFWLRVPKVA-----SHLEGPSLHWSYGLRP 49
DB 2 HWSYGLRPGSGPSLDEKIA-----KMEKASSVFNVNSSSGPSLHWSYGLRP 50
RESULT 8
ABR82479
ID ABR82479 standard; protein; 708 AA.
XX ABR82479;
XX 20-NOV-2003 (first entry)
XX Modified human CEA-TT P2 and P30 epitopes.
XX CEA; immune response; carcinoembryonic antigen; antigen presenting cell;
KW APC; cytostatic; vaccine; human; tetanus toxoid; p2; p30; antigen.
XX OS Synthetic.
XX Key Location/Qualifiers
FT Peptide 1..34
FT /note= "signal peptide"
FT Protein 35..708
FT /note= "mature protein"
XX WO2003059379-A2.
XX 24-JUL-2003.
XX 17-JAN-2003; 2003WO-DK000031.
XX PF
XX 17-JAN-2002; 2002DK-00000082.
XX PR
XX 17-JAN-2002; 2002US-0350047F.
XX PHAR-) PHARMEXA AS.
XX Klysner S, Voldborg B;
XX WPI; 2003-587260/55.
XX DR
XX N-PSDB; ACF35966.
XX Inducing an immune response in humans against autologous carcinoembryonic antigen (CEA) comprises administering a modified CEA polypeptide, a nucleic acid encoding the polypeptide, or a microorganism expressing the polypeptide.
XX Disclosure; Page 121-124; 140pp; English.
XX The invention relates to inducing an immune response against autologous carcinoembryonic antigen (CEA) in an animal, e.g. human. The method involves effecting uptake and processing by antigen presenting cells (APCs) in the animal of at least 1 modified CEA polypeptide or of a nucleic acid encoding the modified CEA polypeptide or of a microorganism or virus expressing the modified CEA polypeptide to induce a CTL response and an antibody response that targets the autologous CEA. The method is useful in immunizing actively against diseases characterized by cells that express CEA. The present sequence represents a modified human CEA polypeptide that has tetanus toxoid (TT) P2 and P30 epitopes introduced in its sequence
XX Sequence 708 AA;
QY 11 SSGPS-----LFNNFTVSFWLRVPKVSASHLEGPSLH 42
Query Match 46.0%; Score 136; DB 7; Length 708;
Best Local Similarity 58.4%; Pred. No. 8.4e-08;
Matches 26; Conservative 2; Mismatches 4; Indels 6; Gaps 1;

630 ASNFSPQYSWRIFNFTVSWLRVSKVSASHLEIPQOH 667

RESULT 9

ABR82478
ID ABR82478 standard; protein; 717 AA.

XX ABR82478;

XX 20-NOV-2003 (first entry)

XX Modified human CEA-TT P2 and P30 epitopes.

XX CEA; immune response; carcinoembryonic antigen; antigen presenting cell;
XX APC; cytostatic; vaccine; human; tetanus toxoid; P2; P30; antigen.

XX Synthetic.

XX Key Location/Qualifiers

XX Peptide 1..34 "signal peptide"

XX Protein 35..717

XX /note= "mature protein"

XX WO2003059379-A2.

XX 24-JUL-2003.

XX 17-JAN-2003; 2003WO-DK000031.

XX 17-JAN-2002; 2002DK-00000082.

XX 17-JAN-2002; 2002US-0350047P.

XX (PHAR-) PHARMEXA AS.

XX Klysner S, Voldborg B;

XX WPI; 2003-587260/55.

XX N-PSDB; ACF35964.

XX Inducing an immune response in humans against autologous carcinoembryonic
XX antigen (CEA) comprises administering a modified CEA polypeptide, a
XX nucleic acid encoding the polypeptide, or a microorganism expressing the
XX polypeptide.

XX Disclosure; Page 114-117; 140pp; English.

XX The invention relates to inducing an immune response against autologous
XX carcinoembryonic antigen (CEA) in an animal, e.g. human. The method
XX involves effecting uptake and processing by antigen presenting cells
XX (APCs) in the animal of at least 1 modified CEA polypeptide or of a
XX nucleic acid encoding the modified CEA polypeptide, or of a microorganism
XX or virus expressing the modified CEA polypeptide to induce a CTL response
XX and an antibody response that targets the autologous CEA. The method is
XX useful in immunizing actively against diseases characterized by cells
XX that express CEA. The present sequence represents a modified human CEA
XX polypeptide that has tetanus toxoid (Tt) P2 and P30 epitopes introduced
XX in its sequence

XX Sequence 717 AA;

Query Match 45.6%; Score 125; DB 7; Length 717;

Best Local Similarity 96.0%; Pred. No. 1.2e-07;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 16 LFNNFTVSWLRVSKVSASHLEGPS 40

Db 692 LFNNFTVSWLRVSKVSASHLEGPS 716

RESULT 10

AAB20150

ID XX AAB20150 standard; protein; 109 AA.
AC XX AAB20150;
DT XX 30-APR-2001 (first entry)
DE XX Growth differentiation factor 8 AutoVac construct GDF-8 P30-3A.
XX Growth differentiation factor 8; GDF-8; myostatin; tetanus toxin;
KW T-cell epitope; down-regulation; vaccine; muscle; meat; cachexia;
KW cardiant; human; mutant; mutein.
XX Homo sapiens.
OS Clostridium tetani.
OS Synthetic.
OS Chimeric.
XX Key Location/Qualifiers
FH Region 1..78
FT /note= "identical to residues 267-345 of human GDF-8"
FT Misc-difference 73
FT /note= "Cys-73 may be substituted by Ser to avoid
FT disulfide bond formation"
FT Region 79..99
FT /note= "tetanus toxoid P2 epitope"
FT Misc-difference 90..91
FT /note= "optionally replaced by Glu-Gly"
FT Region 100..109
FT /note= "identical to residues 366-375 of human GDF-8"
XX WO200105820-A2.
PN 25-JAN-2001.
XX 20-JUL-2000; 2000WO-DK000413.
PR 20-JUL-1999; 99DK-00001014.
PR 26-JUL-1999; 99US-0145275P.
XX (MEBI-) M & E BIOTECH AS.
XX Halkier T, Mouritsen S, Klysner S;
XX WPI; 2001-112680/12.
DR Increasing the muscle mass of animals used in meat production by down
PT regulating growth differentiation factor 8 (GDF-8) activity in the animal
PT through induction of anti-GDF-8 antibody production.
XX Example 1; Page 102-103; 110pp; English.
PS The present sequence is that of AutoVac construct GDF-8 P30-3A,
CC comprising the 109 C-terminal amino acid residues of human growth
CC differentiation factor 8 (GDF-8) in which residues 79-99 are replaced by
CC the promiscuous tetanus toxin T-cell epitope P30 (see AAB20144). It is an
CC object of the invention to produce a recombinant therapeutic vaccine that
CC is capable of effecting down-regulation of GDF-8 in order to increase the
CC muscle growth rate of farm animals. The vaccines (see AAB20145-53) are
CC capable of breaking autotolerance against autologous GDF-8. They comprise
CC the C-terminal portion of human GDF-8 in which a portion of the native
CC sequence is replaced by a T-cell epitope such as P30, with minimal
CC disturbance of the authentic 3-dimensional structure of the protein.
CC Nucleic acids encoding the GDF-8 variants can be used for genetic
CC immunization of the animals. Down-regulation of GDF-8 activity can
CC increase muscle mass by up to at least 45% in cattle, pigs and poultry
CC used for meat production, reducing the need for antibiotic feed-
CC additives. Anti-GDF8 vaccines can be used to treat human diseases such as
CC cancer cachexia where muscle atrophy is pronounced and for patients
CC suffering from acute and chronic heart failure
XX Sequence 109 AA;

Query Match

44.9%; Score 123; DB 4; Length 109;

```

Best Local Similarity 78.1%; Pred. No. 2.3e-08;
Matches 25; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 10 GSSG----PSLFNNFTVSFWLRVVKVSASHLE 37
   |||:|||||:|||||:|||||:|||||:|||||:
Db 68 GSAGPCCTPTKFNFTVSFWLRVVKVSASHLE 99

RESULT 11
AAO30488
ID AAO30488 standard; protein; 194 AA.
XX
AC AAO30488;
XX
DT 22-SEP-2003 (first entry)
XX
DE Human TNFalpha variant, TNF34-P2-P30.
XX
KW Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;
KW tumour necrosis factor alpha; gene therapy; arthritis; human; mutant;
KW mutein; variant; tetanus toxoid; epitope.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 2..109
FT /note= "Human TNF"
FT Region 110..124
FT /note= "Tetanus toxoid P2 epitope"
FT Region 125..145
FT /note= "Tetanus toxoid P30 epitope"
FT Region 146..194
FT /note= "Human TNF"
XX
PN WO2003042244-A2.
XX
PD 22-MAY-2003.
XX
PF 15-NOV-2002; 2002WO-DK000764.
XX
PR 16-NOV-2001; 2001DK-00001702.
XX
PR 16-NOV-2001; 2001US-0331575P.
XX
PA (PHAR-) PHARMEXA AS.
PA (KLYS/) KLYSNER S.
PA (NIEL/) NIELSEN F S.
PA (BRAT/) BRATT T.
PA (VOLD/) VOLDORF B.
PA (MOUR/) MOURITSEN S.
XX
PI Klyner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;
XX
DR WPI; 2003-449558/42.
XX
PT New immunogenic analogue of a polymeric protein, useful for preparing a
PT composition for treating inflammatory diseases e.g. arthritis.
XX
PS Claim 23; Page 158; 196pp; English.
XX
CC The invention relates to immunogenic analogues of multimeric proteins
CC such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis
CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic
CC analogues. The immunogenic analogue is useful for preparing a composition
CC for treating inflammatory diseases, e.g., arthritis. It is also used in
CC gene therapy. The present sequence is human TNFalpha variant protein with
CC an inserted tetanus toxoid P2 and P30 epitopes. This sequence is used to
CC illustrate the method of the invention
XX
SQ Sequence 194 AA;
Query Match 44.9%; Score 123; DB 6; Length 194;

Best Local Similarity 82.1%; Pred. No. 4.6e-08;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 16 LFNFTVSFWLRVVKVSASHLEGP SLHW 43
   |||:|||||:|||||:|||||:|||||:
Db 124 LFNFTVSFWLRVVKVSASHLEAEAKPW 151

RESULT 12
AA92633
ID AAY92633 standard; protein; 750 AA.
XX
AC AAY92633;
XX
DT 10-AUG-2000 (first entry)
XX
DE Mutant human prostate specific membrane antigen construct, hPSM1.10.
XX
KW Prostate specific membrane antigen; immunogenized construct; mutant;
KW vaccination; cytotoxic T-lymphocyte immunity; breast cancer;
KW prostate cancer; cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 24..38
FT /label= P2
FT /note= "foreign epitope"
FT Peptide 673..693
FT /label= P30
FT /note= "foreign epitope"
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK000525.
XX
PR 05-OCT-1998; 98DK-00001261.
XX
PR 20-OCT-1998; 98US-0105011P.
XX
PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated peptide
PT antigens for the treatment of breast and prostate cancer.
XX
PS Example 1; Page; 220pp; English.
XX
CC AAY92627-49 are mutant immunogenized human prostate specific membrane
CC antigen (PSM) constructs, which contain foreign epitopes (P2 and/or P30).
CC The immunogenic analogues of PSM can be used in the claimed method as an
CC autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody
CC binding regions and cysteine residues involved in disulfide bonds are
CC preserved in the immunogenized forms. The method is used for inducing
CC immune responses against weakly immunogenic cell-associated peptide
CC antigens (PA) such as those associated with cancers (self-proteins), e.g.
CC human prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human

```

CC PSM, FGF8b and Her2, respectively. Note: This sequence was constructed from the wild type human PSM (AA92619), which appears on pages 184-187 of the specification

XX Sequence 750 AA;

Query Match 44.3%; Score 121.5; DB 3; Length 750;

Best Local Similarity 86.2%; Pred. No. 3.5e-07; Indels 1; Gaps 1;

Matches 25; Conservative 0; Mismatches 3;

QY 17 FNNFTVFWLRVFKVSASHLEGPSSLSWSY 45

DB 673 FNNFTVFWLRVFKVSASHLE-PSSHNY 700

RESULT 13

ID AA92665 standard; peptide; 216 AA.

XX AC AA92665;

DT 10-AUG-2000 (first entry)

XX MUC-1 analogue containing foreign epitopes.

XX Mucin repeat; MUC-1 analogue; vaccination; self-protein; cancer;

KW cytotoxic T-lymphocyte immunity; breast cancer; prostate cancer;

KW cell-associated peptide antigen; foreign epitope.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 61..75

FT /label= P2

FT Peptide 136..156

FT /label= P30

FT /note= "q"

XX WO200020027-A2.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-DK000525.

XX 05-OCT-1998; 98DK-00001261.

XX 20-OCT-1998; 98US-0105011P.

XX (MEBI-) M & E BIOTECH AS.

XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;

PI Gautam A, Birk P, Karlsson G;

XX WPI; 2000-349917/30.

XX Inducing immune responses to weakly immunogenic, tumor associated peptide

XX antigens for the treatment of breast and prostate cancer.

XX Example 4; Page; 220pp; English.

XX This is an immunogenized MUC-1 analogue containing foreign epitopes P2

XX and P30. Immunogenic analogues of MUC-1 and, e.g. human prostate specific

XX membrane antigen (hPSM) can be used in the claimed method as an

XX autovaccine to induce a CTL response. Subdominant CTL epitopes, antibody

XX binding regions and cysteine residues involved in disulfide bonds are

XX preserved in the immunogenized forms (see features table). 10 regions

XX suitable for the insertion of foreign T helper epitopes were identified.

XX The method is used for inducing immune responses against weakly

XX immunogenic cell-associated peptide antigens (PA) such as those

XX associated with cancers (self-proteins), e.g. hPSM, heregulin 2 (Her2)

XX and/or fibroblast growth factor 8b (FGF8b). The method comprises

XX effecting simultaneous presentation by antigen producing cells (APCs) of

CC cell-associated PA; and (2) at least 1 first T helper cell group which is

CC foreign to the animal. Analogues of human PSM, human Her2 and

CC human/murine FGF8b comprising a substantial part of all known and

CC predicted CTL and B-cell epitopes of the respective PA and including at

CC least one foreign T helper epitope are also claimed. The method is used

CC to treat prostate, prostate/breast or breast cancer when the PA is human

CC PSM, FGF8b and Her2, respectively. Note: This sequence does not appear in

CC the specification. It was made using the mucin repeat sequence

CC (AA92664), P2 and P30 (AA92625-26), which appear on pages 220, 213 and

CC 214 respectively, of the specification

XX Sequence 216 AA;

Query Match 44.2%; Score 121; DB 3; Length 216;

Best Local Similarity 64.1%; Pred. No. 9.6e-08;

Matches 25; Conservative 2; Mismatches 2; Indels 10; Gaps 1;

QY 9 PGSSGP-----SLFNNFTVFWLRVFKVSASHLE 37

DB 118 PGSTAPPAGVTSAPDTRFNNFTVFWLRVFKVSASHLE 156

RESULT 14

AAO30459

ID AAO30459 standard; protein; 287 AA.

XX AC AAO30459;

XX 22-SEP-2003 (first entry)

XX hIL5.36 variant protein.

DE hIL5.36

KW Multimeric protein; interleukin 5; IL5; TNFalpha; inflammatory disease;

KW tumour necrosis factor alpha; gene therapy; arthritis; interleukin 5;

KW IL5; epitope; human; tetanus toxoid; chimeric; variant; mutant; mutein.

XX Homo sapiens.

XX Unidentified.

XX Chimeric.

XX Key Location/Qualifiers

FT Peptide 1..19

FT /note= "Human IL5 leader peptide"

FT Protein 20..287

FT /note= "Mature hIL5.36 protein"

FT Region 24..44

FT /note= "Tetanus toxoid P30 epitope"

FT Region 273..287

FT /note= "Tetanus toxoid P2 epitope"

XX WO2003042244-A2.

XX 22-MAY-2003.

XX 15-NOV-2002; 2002WO-DK000764.

XX 16-NOV-2001; 2001DK-00001702.

XX 16-NOV-2001; 2001US-0331575P.

XX (PHAR-) PHARMEXA AS.

XX (KLYS-) KLYSNER S.

XX (NIEL-) NIELSEN F S.

XX (BRAT-) BRATT T.

XX (VOLD-) VOLDBOG B.

XX (MOUR-) MOURITSEN S.

XX Klysner S, Nielsen FS, Bratt T, Voldborg B, Mouritsen S;

XX WPI; 2003-449558/42.

XX N-PSDB; AAL61295.

XX New immunogenic analogue of a polymeric protein, useful for preparing a

XX composition for treating inflammatory diseases e.g. arthritis.

XX Claim 20; Page 115-117; 196pp; English.

PS The invention relates to immunogenic analogues of multimeric proteins

XX such as immunogenic variants of interleukin 5 (IL5) and tumour necrosis

CC factor alpha (TNF, TNFalpha) and methods for production of immunogenic

CC analogues. The immunogenic analogue is useful for preparing a composition

CC for treating inflammatory diseases, e.g., arthritis. It is also used in

CC gene therapy. The present sequence is a fusion construct variant which

CC comprises 2 human interleukin 5 (IL5) monomers joined by diglycine linker

CC and including terminally positioned tetanus toxoid epitopes P30 and P2.

CC This sequence is used to illustrate the method of the invention

XX SQ Sequence 287 AA;

Query Match 44.2%; Score 121; DB 6; Length 287;

Best Local Similarity 85.2%; Pred. No. 1.3e-07;

Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 14 PSLFNNFTVSWLRVPKVSASHLEGPS 40

DB 21 PTEFNNFTVSEWLRVPKVSASHLEIPT 47

RESULT 15

ABR82481

ID ABR82481 standard; protein; 537 AA.

XX AC ABR82481;

XX DT 20-NOV-2003 (first entry)

XX DE Truncated human CEA-TT P2 and P30 epitopes.

XX KW CEA; immune response; carcinoembryonic antigen; antigen presenting cell;

XX KW APC; cytostatic; vaccine; human; tetanus toxoid; P2; P30; antigen.

XX OS Synthetic.

XX FH Key

FT Peptide 1..34

FT Protein /note= "signal peptide"

FT /note= "mature protein"

XX WO2003059379-A2.

XX PD 24-JUL-2003.

XX PF 17-JAN-2003; 2003WO-DK000031.

XX PR 17-JAN-2002; 2002DK-00000082.

XX PR 17-JAN-2002; 2002US-0350047P.

XX PA (PHAR-) PHARMEXA AS.

XX PI Klysner S, Voldborg B;

XX DR WPI; 2003-587260/55.

XX DR N-PSDB; ACF35968.

XX PT Inducing an immune response in humans against autologous carcinoembryonic

FT antigen (CEA) comprises administering a modified CEA polypeptide, a

FT nucleic acid encoding the polypeptide, or a microorganism expressing the

FT polypeptide.

XX PS Disclosure; Page 134-137; 14Opp; English.

XX CC The invention relates to inducing an immune response against autologous

CC carcinoembryonic antigen (CEA) in an animal, e.g. human. The method

CC involves effecting uptake and processing by antigen presenting cells

CC (APCs) in the animal of at least 1 modified CEA polypeptide or of a

CC nucleic acid encoding the modified CEA polypeptide or of a microorganism

CC or virus expressing the modified CEA polypeptide to induce a CTL response

CC and an antibody response that targets the autologous CEA. The method is

CC useful in immunizing actively against diseases characterized by cells

CC that express CEA. The present sequence represents a truncated human CEA

CC polypeptide that has tetanus toxoid (TT) P2 and P30 epitopes introduced

CC in its sequence

XX SQ Sequence 537 AA;

Query Match 44.2%; Score 121; DB 7; Length 537;

Best Local Similarity 95.8%; Pred. No. 2.8e-07;

Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 17 FNNFTVSWLRVPKVSASHLEGPS 40

DB 513 FNNFTVSWLRVPKVSASHLEGTS 536

Search completed: March 10, 2004, 09:12:13

Job time : 75.2918 secs

```

US-09-848-834A-18
Sequence 18, Application US/09848834A
Patent No. US20020076416A1
GENERAL INFORMATION:
APPLICANT: Aphton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIOR FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 18
LENGTH: 50
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chimeric peptide consisting
OTHER INFORMATION: GNRH linked by a spacer to
OTHER INFORMATION: anus toxoid precursor (Tent
OTHER INFORMATION: o amino acid sequence 2-10
NAME/KEY: MOD RES
LOCATION: (1)..(1)
OTHER INFORMATION: Pyroglutamic acid or 5-oxop
NAME/KEY: MOD RES
LOCATION: (50)..(50)
OTHER INFORMATION: Amidated glycine or glycina
NAME/KEY: PEPTIDE
LOCATION: (1)..(10)
OTHER INFORMATION: Amino acid sequence 1-10 of
NAME/KEY: PEPTIDE
LOCATION: (1)..(16)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (17)..(37)
OTHER INFORMATION: Amino acid sequence 947-967

```

OTHER INFORMATION: oxylysine
NAME/KEY: PEPTIDE
LOCATION: (38)..(41)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (42)..(50)
OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-18

Query Match 99.3%; Score 272; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-26;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGFSLNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49
Db 2 HWSYGLRPGSSGFSLNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49

RESULT 2

US-09-848-834A-14
Sequence 14, Application US/09848834A
Patent No. US20020076416A1

GENERAL INFORMATION:
APPLICANT: Aphton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047
CURRENT APPLICATION NUMBER: US/09/848,834A
PRIORITY FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIORITY FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 37

TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the
OTHER INFORMATION: rRH hormone linked by a spacer to amino acid sequence 947-967 of
OTHER INFORMATION: the Tetanus toxoid precursor (Tentoxylisin)

NAME/KEY: MOD_RES
LOCATION: (1)..(1)

OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: PEPTIDE
LOCATION: (11)..(16)
OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE

OTHER INFORMATION: Spacer peptide

NAME/KEY: PEPTIDE

LOCATION: (17)..(37)

OTHER INFORMATION: (Tentoxylisin)

US-09-848-834A-14

Query Match 72.6%; Score 199; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSGFSLNNFTVSFWLRVPKVSASHLE 37
Db 2 HWSYGLRPGSSGFSLNNFTVSFWLRVPKVSASHLE 37

RESULT 3

US-09-848-834A-10
Sequence 10, Application US/09848834A
Patent No. US20020076416A1

GENERAL INFORMATION:
APPLICANT: Aphton Corporation
TITLE OF INVENTION: Chimeric Peptide Immunogens
FILE REFERENCE: 1102865-0047

CURRENT APPLICATION NUMBER: US/09/848,834A

CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/202,328
PRIORITY FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 34
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 947-967 of the
OTHER INFORMATION: Tetanus toxoid precursor (Tentoxylisin) linked by a spacer to an
OTHER INFORMATION: amino acid sequence 2-10 of the GnRH hormone
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: Amidated phenylalanine
NAME/KEY: PEPTIDE
LOCATION: (1)..(21)
OTHER INFORMATION: Amino acids 947-967 of the Tetanus Toxoid Precursor
NAME/KEY: PEPTIDE
LOCATION: (22)..(25)
OTHER INFORMATION: Spacer peptide
NAME/KEY: PEPTIDE
LOCATION: (26)..(34)
OTHER INFORMATION: Amino acids 2-10 of the human GnRH hormone
NAME/KEY: MOD_RES
LOCATION: (34)..(34)
OTHER INFORMATION: Amidated glycine or glycineamide
US-09-848-834A-10

Query Match 67.5%; Score 185; DB 9; Length 34;
Best Local Similarity 100.0%; Pred. No. 9.5e-16;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 49
Db 1 FNNFTVSFWLRVPKVSASHLEGPSLHWSYGLRP 33

RESULT 4

US-09-848-834A-19
Sequence 19, Application US/09848834A
Patent No. US20020076416A1

GENERAL INFORMATION:

APPLICANT: Aphton Corporation

TITLE OF INVENTION: Chimeric Peptide Immunogens

FILE REFERENCE: 1102865-0047

CURRENT APPLICATION NUMBER: US/09/848,834A

PRIORITY FILING DATE: 2001-05-04

PRIOR APPLICATION NUMBER: 60/202,328

PRIORITY FILING DATE: 2000-05-05

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn version 3.0

SEQ ID NO 19

LENGTH: 46

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of human
OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 830-844 of Tetan
OTHER INFORMATION: toxoid precursor (Tentoxylisin) linked by a spacer to amino aci
OTHER INFORMATION: sequence 1-10 of GnRH

NAME/KEY: MOD_RES

LOCATION: (1)..(1)

OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline

NAME/KEY: MOD_RES

LOCATION: (46)..(46)

OTHER INFORMATION: Amidated glycine or glycineamide

NAME/KEY: PEPTIDE

LOCATION: (1)..(10)

OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone

NAME/KEY: PEPTIDE

/ LOCATION: (11)..(16)
/ OTHER INFORMATION: Spacer peptide
/ NAME/KEY: PEPTIDE
/ LOCATION: (17)..(31)
/ OTHER INFORMATION: Amino acid sequence 830-844 of the Tetanus toxoid precursor
/ NAME/KEY: PEPTIDE (Tentoxylisin)
/ LOCATION: (32)..(37)
/ OTHER INFORMATION: Spacer peptide
/ NAME/KEY: PEPTIDE
/ LOCATION: (38)..(46)
/ OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-19

Query Match 52.9%; Score 145; DB 9; Length 46;
Best Local Similarity 60.4%; Pred. No. 1.1e-10;
Matches 29; Conservative 6; Mismatches 9; Indels 4; Gaps 1;
Qy 2 HWSYGLRPGSGPSLNNFTVSWLRVVKVSAHLSHLEGPGLHWSYGLRP 49
Db 2 HWSYGLRPGSGPSLQYIKANSKFIGITELSS---GPSLHWSYGLRP 45

RESULT 5
US-09-848-834A-17
/ Sequence 17, Application US/09848834A
/ Patent No. US20020076416A1
/ GENERAL INFORMATION:
/ APPLICANT: Apton Corporation
/ TITLE OF INVENTION: Chimeric Peptide Immunogens
/ CURRENT APPLICATION NUMBER: US/09/848,834A
/ PRIOR FILING DATE: 2001-05-04
/ PRIOR FILING DATE: 2002-05-05
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 17
/ LENGTH: 47
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of the Gn
/ OTHER INFORMATION: RH hormone linked by a spacer to amino acid sequence 288-302 of
/ OTHER INFORMATION: the Measles virus protein F linked by a spacer to amino acid seq
/ OTHER INFORMATION: uence 2-10 of the GnRH hormone
/ NAME/KEY: MOD RES
/ LOCATION: (11)..(11)
/ OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
/ NAME/KEY: MOD RES
/ LOCATION: (47)..(47)
/ OTHER INFORMATION: Amidated-glycine or glycylamide
/ NAME/KEY: PEPTIDE
/ LOCATION: (1)..(10)
/ OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
/ NAME/KEY: PEPTIDE
/ LOCATION: (11)..(18)
/ OTHER INFORMATION: Spacer peptide
/ NAME/KEY: PEPTIDE
/ LOCATION: (19)..(34)
/ OTHER INFORMATION: Amino acid sequence 288-302 of the Measles virus fusion protein,
/ NAME/KEY: PEPTIDE
/ LOCATION: (35)..(38)
/ OTHER INFORMATION: Spacer peptide
/ NAME/KEY: PEPTIDE
/ LOCATION: (39)..(47)
/ OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-17

Query Match 51.6%; Score 141.5; DB 9; Length 47;
Best Local Similarity 59.8%; Pred. No. 3e-10;
Matches 30; Conservative 1; Mismatches 11; Indels 9; Gaps 2;

Qy 2 HWSYGLRPGSGPSLNNFTVSWLRVVKVSAHLSHLEGPGLHWSYGLRP 49
Db 2 HWSYGLRPGSGPSL-----KLLSEIKGVIVHRLEGVEGFSLHWSYGLRP 46

RESULT 6
US-09-848-834A-20
/ Sequence 20, Application US/09848934A
/ Patent No. US20020076416A1
/ GENERAL INFORMATION:
/ APPLICANT: Apton Corporation
/ TITLE OF INVENTION: Chimeric Peptide Immunogens
/ FILE REFERENCE: 1102865-0047
/ CURRENT APPLICATION NUMBER: US/09/848,834A
/ PRIOR FILING DATE: 2001-05-04
/ PRIOR FILING DATE: 2000-05-05
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 20
/ LENGTH: 51
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Chimeric peptide consisting of amino acid sequence 1-10 of huma
/ OTHER INFORMATION: GnRH linked by a spacer to amino acid sequence 378-398 of Plas
/ OTHER INFORMATION: ium falciparum circumsporozoite (CSP) protein
/ NAME/KEY: MOD RES
/ LOCATION: (1)..(1)
/ OTHER INFORMATION: Pyroglutamic acid or 5-oxoproline
/ NAME/KEY: MOD RES
/ LOCATION: (51)..(51)
/ OTHER INFORMATION: Amidated glycine or glycylamide
/ NAME/KEY: PEPTIDE
/ LOCATION: (1)..(10)
/ OTHER INFORMATION: Amino acid sequence 1-10 of the human GnRH hormone
/ NAME/KEY: PEPTIDE
/ LOCATION: (11)..(16)
/ OTHER INFORMATION: Spacer peptide
/ NAME/KEY: PEPTIDE
/ LOCATION: (17)..(36)
/ OTHER INFORMATION: Amino acid sequence 378-398 of the Plasmodium falciparum
/ OTHER INFORMATION: circumsporozoite (CSP) protein
/ NAME/KEY: PEPTIDE
/ LOCATION: (37)..(42)
/ OTHER INFORMATION: Spacer peptide
/ NAME/KEY: PEPTIDE
/ LOCATION: (43)..(51)
/ OTHER INFORMATION: Amino acid sequence 2-10 of the human GnRH hormone
US-09-848-834A-20

Query Match 50.5%; Score 138.5; DB 9; Length 51;
Best Local Similarity 54.7%; Pred. No. 7.8e-10;
Matches 29; Conservative 4; Mismatches 11; Indels 9; Gaps 2;
Qy 2 HWSYGLRPGSGPSLNNFTVSWLRVVKVSA-----SHLEGPGLHWSYGLRP 49
Db 2 HWSYGLRPGSGPSLDEKXIA----KMEKASSVFNWVNSSGSPSLHWSYGLRP 50

RESULT 7
US-10-295-074-46
/ Sequence 46, Application US/10295074
/ Publication No. US20030185845A1
/ GENERAL INFORMATION:
/ APPLICANT: Pharmexa A/S
/ TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
/ FILE REFERENCE: P1013DK00
/ CURRENT APPLICATION NUMBER: US/10/295,074
/ CURRENT FILING DATE: 2002-11-15
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 46

LENGTH: 194
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: hTNF with inserted tetanus toxoid P2 and P30 epitopes
NAME/KEY: MUTAGEN
LOCATION: (110)..(124)
OTHER INFORMATION: Tetanus toxoid P2 epitope (SEQ ID NO: 2)
FEATURE:
NAME/KEY: MUTAGEN
LOCATION: (125)..(145)
OTHER INFORMATION: Tetanus toxoid P30 epitope (SEQ ID NO: 3)
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (2)..(109)
OTHER INFORMATION: hTNF amino acids 1-108
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (146)..(194)
OTHER INFORMATION: hTNF amino acids 109-157
US-10-295-074-46

Query Match 44.9%; Score 123; DB 14; Length 194;
Best Local Similarity 82.1%; Pred. No. 2.7e-07;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 16 LFNFTVSFWLRVPKVSASHLEGPSLHW 43
DB 124 LFNFTVSFWLRVPKVSASHLEAKPW 151

RESULT 8
US-10-295-074-13
Sequence 13, Application US/10295074
Publication No. US20030185845A1
GENERAL INFORMATION:
APPLICANT: Pharmexa A/S
TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
FILE REFERENCE: P1013DK00
CURRENT APPLICATION NUMBER: US/10/295, 074
CURRENT FILING DATE: 2002-11-15
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 287
TYPE: PRT
FEATURE:
ORGANISM: Artificial sequence
OTHER INFORMATION: Two human IL5 monomers joined by diglycine linker and including
OTHER INFORMATION: terminally positioned P30 and P2 epitopes
US-10-295-074-13

Query Match 44.2%; Score 121; DB 14; Length 287;
Best Local Similarity 85.2%; Pred. No. 7.5e-07;
Matches 23; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 14 PSLFNFTVSFWLRVPKVSASHLEGPS 40
DB 21 PTFNFTVSFWLRVPKVSASHLEIPT 47

RESULT 9
US-10-295-074-11
Sequence 11, Application US/10295074
Publication No. US20030185845A1
GENERAL INFORMATION:
APPLICANT: Pharmexa A/S
TITLE OF INVENTION: NOVEL IMMUNOGENIC MIMETICS OF MULTIMER PROTEINS
FILE REFERENCE: P1013DK00
CURRENT APPLICATION NUMBER: US/10/295, 074
CURRENT FILING DATE: 2002-11-15
NUMBER OF SEQ ID NOS: 60

SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 285
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: 2 human IL5 monomers joined by P2 and P30 epitopes
US-10-295-074-11

Query Match 43.8%; Score 120; DB 14; Length 285;
Best Local Similarity 92.0%; Pred. No. 9.8e-07;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 16 LFNFTVSFWLRVPKVSASHLEGPS 40
DB 149 LFNFTVSFWLRVPKVSASHLEIPT 173

RESULT 10
US-10-297-942-18
Sequence 18, Application US/10297942
Publication No. US20030185816A1
GENERAL INFORMATION:
APPLICANT: Ferring BV
TITLE OF INVENTION: Solubilised Protein Vaccines
FILE REFERENCE: P68445US0
CURRENT APPLICATION NUMBER: US/10/297, 942
CURRENT FILING DATE: 2003-04-21
PRIOR APPLICATION NUMBER: PCT/DK01/00431
PRIOR FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: DK PA 2000 00966
PRIOR FILING DATE: 2000-06-21
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 18
LENGTH: 158
TYPE: PRT
ORGANISM: Homo sapiens
US-10-297-942-18

Query Match 43.2%; Score 118.5; DB 14; Length 158;
Best Local Similarity 80.6%; Pred. No. 7.8e-07;
Matches 25; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 7 LRPGSSGSPSLFNFTVSFWLRVPKVSASHLE 37
DB 2 VRSSRTPS-FNNFTVSFWLRVPKVSASHLE 31

RESULT 11
US-10-297-942-12
Sequence 12, Application US/10297942
Publication No. US20030185816A1
GENERAL INFORMATION:
APPLICANT: Ferring BV
TITLE OF INVENTION: Solubilised Protein Vaccines
FILE REFERENCE: P68445US0
CURRENT APPLICATION NUMBER: US/10/297, 942
CURRENT FILING DATE: 2003-04-21
PRIOR APPLICATION NUMBER: PCT/DK01/00431
PRIOR FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: DK PA 2000 00966
PRIOR FILING DATE: 2000-06-21
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 158
TYPE: PRT
ORGANISM: Homo sapiens
US-10-297-942-12

Query Match 43.1%; Score 118; DB 14; Length 158;
Best Local Similarity 100.0%; Pred. No. 9e-07;